Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

DELOS D. WICKENS, Associate Editor
Ohio State University

Consulting Editors

E. JAMES ARCHER, *University of Wisconsin*
JUDSON S. BROWN, *State University of Iowa*
CLETUS J. BURKE, *Indiana University*
PAUL M. FITTS, *Ohio State University*
FREDERICK C. FRICK, *Massachusetts Institute of Technology*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
LLOYD G. HUMPHREYS, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*

CONTENTS

Ready Signals and the Effect of Interpolated UCS Presentations in Eyelid Conditioning: R. H. DUFORT AND G. A. KIMBLE 1
Differential Effects of "Display Lags" and "Control Lags" on the Performance of Manual Tracking Systems: W. D. GARVEY, J. S. SWEENEY, AND H. P. BIRMINGHAM 8
Influence of Intertrial Interval During Extinction on Spontaneous Recovery of Conditioned Eyelid Responses: M. G. HOWAT AND D. A. GRANT 11
Independence in the Learning of Two Consecutive Responses Per Trial: A. L. BRODY 16
Discrimination Learning in a Verbal Conditioning Situation: J. POPPER AND R. C. ATKINSON 21
Effect of the Spatial Relationship Between Cue, Reward, and Response in Simple Discrimination Learning: J. V. MURPHY AND R. E. MILLER 26
Delay of Reinforcement, Response Perseveration, and Discrimination Reversal: B. H. PUBOLS, JR. 32
Discrimination Learning as a Function of Pretraining Reinforcement Schedules: H. W. STEVENSON AND L. A. PIROJNIKOFF 41
A Positive Correlation Between Action Decrement and Learning: E. L. WALKER AND N. PARADISE 45
Amount and Percentage of Reinforcement and Duration of Goal Confinement in Conditioning and Extinction: S. H. HULSE, JR. 48
Complex Incidental Learning as a Function of Anxiety and Task Difficulty: C. D. SPIELBERGER, L. D. GOODSTEIN, AND W. G. DAHLSTROM 58
Associative Connection Between Paired Verbal Items: J. RICHARDSON AND A. ERLEBACHER ... 62
Changes in Distribution of Muscular Tension During Psychomotor Performance: L. W. GREGG 70
A Relationship Between Incentive Motivation and Ability Level in Psychomotor Performance: E. A. FLEISHMAN 78
Instrumental and Competing Behavior as a Function of Trials and Reward Magnitude: A. C. PERKBOOM AND B. M. CRAWFORD 82
Studies in Incidental Learning: VII. Effects of Frequency of Exercise and Length of List: L. POSTMAN AND P. A. ADAMS 86
Supplementary Report: Discrimination Learning in Rats as a Function of Highly Distributed Trials: D. F. PENNINGTON AND R. THOMPSON 94
Replication Report: Two- and Three-Choice Verbal-Conditioning Phenomena: J. W. COTTON AND A. RECHTSCHAFFEN 96

American Psychological Association

Vol. 56, No. 1 July 1958

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Circulation Manager
SADIE J. DOYLE, Editorial Assistant

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, 53, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association.*** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1958 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. and R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1958 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 56, NO. 1 JULY, 1958

READY SIGNALS AND THE EFFECT OF INTERPOLATED UCS PRESENTATIONS IN EYELID CONDITIONING[1]

ROBERT H. DUFORT[2] AND GREGORY A. KIMBLE

Yale University *Duke University*

Kimble, Mann, and Dufort (**10**) have shown that, under certain conditions, interpolated UCS trials have the same effect upon eyelid conditioning as trials on which the CS and UCS are presented together. Specifically, after 20 ordinary conditioning trials, these investigators omitted the CS on Trials 21–40 and presented only the UCS. They found that *Ss* conditioned by this method (experimental group) performed at the same level on a final series of 20 conditioning trials (Trials 41–60) as a control group for which all 60 trials consisted of CS-UCS pairings. Pavlov (**14,** p. 139), incidentally, reports a similar phenomenon for salivary conditioning.

Recently, Goodrich, Ross, and Wagner (**3**) tried to obtain this result and failed. On Trials 41–60, *Ss* in their experimental group gave significantly fewer CR's than *Ss* in their control group. Attempting to explain this result, they considered several ways in which the procedures of the two experiments differed without, however, identifying any specific variable(s) as responsible for the contrasting results. The differences they mentioned, with their procedure first, were: (*a*) delayed vs. trace conditioning procedures, (*b*) ready signal before each trial vs. no ready signal, (*c*) irregular vs. regular intertrial intervals, and (*d*) graphic vs. all-or-none recording of CR's. Other differences which might have been mentioned were: (*e*) instructions to *Ss*,[3] (*f*) number and patterning of preliminary trials, (*g*) illumination of the experimental room, (*h*) the ampli-

[1] This research received financial support from the Duke University Research Council, the National Science Foundation (Grant 2147), and the United States Public Health Service (Grant M 1509). In the replication conducted at Duke University, considerable assistance in setting up the equipment and in running *Ss* came from Nora Alston, C. A. Boneau, Nancy Bowles, and Margaret S. King. The assistance of Frank Logan in obtaining the apparatus at Yale University is also acknowledged with thanks.

[2] Now at the University of Richmond.

[3] The USPHS grant (see footnote 1) made it possible for the second author to visit the State University of Iowa laboratory and to obtain first-hand experience with their procedures. This experience was an aid in setting up the experiment. It also provided the basis for several statements in this paper which imply a closer acquaintance with their method than would be possible from the Goodrich, Ross, and Wagner report. We wish to thank these investigators and K. W. Spence for their courtesy.

tude of the response counted as a CR, and (*i*) the intensity of the UCS. In view of the number of differences between the two studies, the best way to identify the important factor would seem to be empirically, by means of an experiment in which the manipulation of some variable produces the Goodrich, Ross, and Wagner (GRW) result under one set of circumstances and the Kimble, Mann, and Dufort (KMD) result under another. The experiment reported here does this, the variable manipulated being the presence or absence of a ready signal.

In the GRW study a ready signal occurred before each trial, a procedure regularly employed in the Iowa laboratory (e.g., **15, 17**), but not in others. Among the important studies in which a ready signal was not used were those of Passey (**13**) and the entire series by Grant and his colleagues (e.g., **4, 5**). In the KMD study, a ready signal was not used, because of the possibility that such a signal might interfere with conditioning. If this were the case, such an effect might obscure the increase in response level produced by interpolated UCS trials, and could account for the failure of the GRW study to reproduce the KMD results. The experiment reported here is a test of this hypothesis. The original KMD study was repeated with and without a ready signal, using apparatus and procedures similar to those of Goodrich, Ross, and Wagner. The prediction was that, with the ready signal, the GRW results would be obtained and that, without it, the KMD results would reappear.

Method

The experiment was performed in two replications, one at Duke University and the other at Yale University. The two replications differed with respect to many incidental details, but were identical as regards the important experimental manipulations. This section presents a description of the Yale replication, followed by a brief statement of the ways in which the Duke replication was different.

Subjects.—The *S*s were 70 paid student nurses who volunteered for the experiment. The records of 6 *S*s could not be scored and were discarded. Four other *S*s were discarded, two because of *E* error and two for not following instructions. Of the remaining 60 *S*s, 15 were run in each of the four experimental groups.

Apparatus.—The *S* sat in an adjustable dental chair in a sound-deadened ventilated chamber, located in one corner of a larger laboratory room containing the apparatus and recording instruments. The wall of the chamber facing *S* was covered with black cloth. A square black box containing the neon CS bulbs was mounted on this wall and partially covered by the cloth. The CS was an increase in brightness from a between-trials level of .15 apparent ft.-candles to 1.32 apparent ft.-candles of a circular milk-glass disc, 2.25 in. in diameter, 43 in. from *S*. The CS disc was shielded on all sides from the chamber illumination (.32 ft.-candles) by four pieces of cardboard, each 2.25 in. long, attached to the box. A constantly running fan in the chamber provided additional soundproofing. Communication with *S* was accomplished by means of an intercommunication system through which *E* could always monitor the chamber while *S* normally could hear nothing from outside.

The duration of the CS on each trial was 550 msec. The UCS was a 3-p.s.i. air puff, limited to a duration of 50 msec. by a 110-v., A.C. solenoid valve, delivered to the corner of the right eye. The UCS was provided by a variable-pressure reducing valve and gauge placed in the building air supply line. The CS and UCS terminated simultaneously, thus providing a CS-UCS interval of 500 msec. Eyelid responses in the range of 200–500 msec. following CS onset were counted as CR's. Sequence and duration of stimulus presentations on each trial were controlled by Hunter electronic timers.

Eyelid responses were recorded by a system essentially the same as that used by Goodrich, Ross, and Wagner, consisting of a microtorque potentiometer mechanically connected to *S*'s right eyelid, a Hunter Eyeblink Amplifier, and an Edin penmotor. This system gave a continuous graphic record, before, during, and after each trial, of the position and movement of the eyelid. An additional pen recorded the onset and termination of CS and UCS.

Experimental design and procedure.—The experiment took the form of a 2 × 2 factorial

design, with ready-signal and no-ready-signal conditions defining one marginal, and control and experimental conditions defining the other. On arrival, *S* was adjusted to the chair and given the instructions, which were the same for all groups except that *Ss* in the two ready-signal conditions were informed of the ready signal and instructed to blink once upon receiving it. All *Ss* then received 5 test trials with the CS alone. Following this, they received 20 conditioning trials on which the CS and UCS were paired. At this point, *Ss* in the experimental groups were given 20 presentations of the UCS alone, while *Ss* in the control groups received 20 more CS-UCS conditioning trials. In assigning *Ss*, an effort was made to match experimental and control groups for response distributions on Trials 1–20. During the UCS trials for the experimental groups, the CS disc remained at the between-trials brightness. All *Ss* were then given a final series of 20 conditioning trials, to make a total of 60 trials. Intertrial intervals of 15, 20, and 25 sec., averaging 20 sec. and randomly arranged in advance, were used.

For one experimental and one control group, a ready signal of 1-sec. duration was presented before each trial. It was provided by a 110-v., A.C. relay wired to produce a buzzing sound when activated. The interval between ready-signal onset and CS onset on test trials and conditioning trials was 3, 4, or 5 sec., randomly varied and arranged in advance. On the UCS trials for the experimental group, the ready signal was presented according to the same time schedule, although the CS was omitted.

At the close of the experimental session, *S* filled out a questionnaire which asked, among other things, about the purpose of the experiment. None of the *Ss* reported recognition of its purpose.

Duke University replication.—This replication was the same as the Yale University replication with respect to all important features except the following: (*a*) *Subjects.* The *Ss* were 33 men and 27 women. All except 6 paid *Ss* were volunteers completing a course requirement. The assignment to experimental and control groups, within conditions, was on the basis of responding on Trials 1–20 without respect to sex for all *Ss* who made at least one CR during Trials 1–20. (*b*) *Apparatus.* Both the CS and UCS were weaker than in the Yale experiment. The CS was the onset of a single, unscreened neon (NE 45) bulb, recessed about .25 in. behind a 3.5 × 8-in. flat black panel, 20 in. directly ahead of *S* and visible through a hole .5 in. in diameter. Its intensity was about .01 apparent ft.-candles. Between trials the CS bulb was dark. The UCS was an air puff of 1.5 p.s.i. In this replication the illumination in the experimental chamber was about 23 apparent ft.-candles, measured at the white wall of the booth. The *S* sat in an office-type chair with his chin in a chin rest looking at the place where the CS appeared. The ready signal was the word "Ready" said by *E* through an intercommunication system, 1–5 sec. before each trial. (*c*) *Procedure.* In this replication the preliminary test trials were two presentations of the CS and three of the UCS, in that order. The intertrial intervals averaged 15 sec., consisting of a regularly repeating sequence of about 11, 11, 16, and 21 sec.

Results

Performance curves for the four groups in each replication appear in Fig. 1. The curves which continue without interruption are those for the control groups receiving 60 conditioning trials; those which are broken between Trials 21 and 40 are for the experimental groups which received only UCS presentations on those trials. As regards the relative positions of the four groups, the two replications quite clearly obtained the same results. The consistently higher level of performance of *Ss* in the Yale replication is probably the result of the stronger UCS used there. For both replications a comparison of the curves in the no-ready-signal condition on the final 20 conditioning trials shows that the groups for which the CS was omitted are performing at the same level as the control groups, confirming the findings of Kimble, Mann, and Dufort. In the ready-signal condition, however, the groups which received UCS presentations are markedly inferior to the control groups on these trials. This result is identical with that reported by Goodrich, Ross, and Wagner.

Statistical analysis supports these observations. In order to evaluate the results of the experiment, the data from the two replications were

combined for analysis of variance. The analysis was performed on both total number of CR's on Trials 41–60 and on gain in CR's over Trials 21–40. The latter measure was obtained by subtracting for each *S* the number of CR's made on Trials 11–20 from the number made on Trials 41–50; it may be taken as a relatively direct index of the effect of the control and experimental procedures during the intervening 20 trials. Table 1 summarizes the analysis. The main effect of replications is highly significant for number of CR's on Trials 41–60, confirming the observation that *S*s in the Yale replication performed at a higher level. When the analysis is applied to gain in CR's over Trials 21–40, however, the replications effect does not reach significance.

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE

Source	*df*	Number of CR's on Trials 41–60		Gain in CR's over Trials 21–40	
		MS	*F*	*MS*	*F*
Control-Exp. (CE)	1	53.34		22.54	
Ready signal—No ready signal (RS)	1	472.04		3.34	
Replications (R)	1	158.70	11.68***	2.70	.55
CE × RS	1	73.62	5.42**	20.82	4.23*
CE × R	1	4.03	.30	.53	.11
RS × R	1	.13	.01	2.13	.43
CE × RS × R	1	3.34	.25	.84	.17
Within	112	13.58		4.92	

* $P = .05$.
** $P = .025$.
*** $P = .001$.

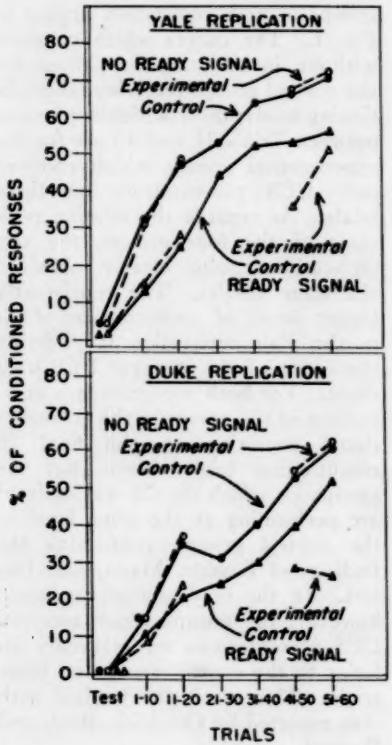

FIG. 1. Conditioning curves for the two replications. Control groups received paired CS-UCS presentations on Trials 21–40. Experimental groups received UCS presentations on Trials 21–40.

With regard to the predicted outcome of the experiment, primary interest is in the interaction of the control-experimental and the ready-signal, no-ready-signal conditions. Table 1 shows that this interaction is significant ($P < .025$) for number of CR's on Trials 41–60. The nature of the interaction was determined by an examination of the simple effects.[4] For the ready-signal condition, the control groups were superior to the experimental groups ($F = 4.10$, $P < .05$); while for the no-ready-signal condition, control and experimental groups did not differ ($F < 1.00$). The ready-signal and

[4] Following Lindquist (11), the simple effects were determined by four individual Treatments × Replications analyses, one for each level of each experimental variable. In none of these was the interaction between the experimental condition and replications significant; the statements which follow, then, appear to hold for both replications.

no-ready-signal control groups did not differ significantly ($F = 2.15$, $P > .10$), while the no-ready experimental groups were superior to the ready-signal experimental groups ($F = 17.66$, $P < .001$).

As Table 1 shows, a significant interaction of control-experimental and ready-signal, no-ready-signal conditions was also obtained on the analysis of gains in CR's over Trials 21–40. Again the simple effects were tested. For the ready-signal condition, the control groups displayed a greater gain than the experimental groups ($F = 7.13$, $P < .001$); while for the no-ready-signal condition, the control and experimental gains were comparable ($F < 1.00$). The gains for ready-signal and no-ready-signal control groups were comparable ($F < 1.00$), while the no-ready-signal experimental groups gained more than the ready-signal experimental groups ($F = 4.61$, $P < .05$).

An important incidental finding is that the level of responding on Trials 1–20 was significantly lower for the ready-signal groups than for the no-ready-signal groups in both Yale ($F = 7.09$, $P < .05$) and Duke ($F = 5.40$, $P < .025$) replications. In no case did the ready signal augment performance in this experiment. Where it had any effect, the effect was an interference.

Discussion

These findings show that the ready signal, as hypothesized, is the factor responsible for the difference between the Goodrich, Ross, and Wagner results and the Kimble, Mann, and Dufort results. The several differences between the two replications reported here contribute to this conclusion; as Underwood has stressed (**18**, p. 286), the fact that similar results have been obtained under somewhat varying circumstances lends weight to the argument that the variable manipulated actually does control the effect produced in the experiment.

The most immediate implication of these results is that they demonstrate an inhibitory effect associated with ready signals. They also support the implication of previous studies in this series (**9, 10**) that associative strength in eyelid conditioning develops to its maximum within the first few conditioning trials[5] and that the protracted rise of the typical eyelid conditioning function chiefly reflects the development of a nonassociative performance factor under the control of the UCS. This statement does not, however, constitute a denial of the continuity theory of conditioning, for there is nothing in such theorizing which requires that the development of associative strength must be slow. In discussing the results of an earlier study, the writers were quite explicit (**9**, p. 390) on the points that (*a*) the data obtained there reflected a continuous although rapid growth of associative strength and (*b*) there were no instances in the data of the step-wise, one-trial conditioning curves reported by Voeks (**19**).

The data obtained in this study do not, of course, demonstrate the nature of the mechanism by which ready signals interfere with performance in eyelid-conditioning studies. Two possibilities suggest themselves. One is that omitting the CS on Trials 21–40 and retaining the ready signal have the same effect as occurs in studies (**1, 12**) which have obtained extinction by lengthening the CS-UCS interval. For, when the CS is omitted, the ready signal must be regarded as having become one, placed

[5] The use of 20 conditioning trials prior to the introduction of the UCS trials was required in order to duplicate this aspect of the Goodrich, Ross, and Wagner study. It does not represent a modification of the Kimble and Dufort (**9**) view that associative strength develops as a more direct function of the number of CR's that *S* has made during the early conditioning trials than of the number of conditioning trials per se.

at a very unfavorable interval with respect to the UCS. The chief problem in such an interpretation is that of explaining how such extinction could interfere with later performance. The most straightforward explanation seems to come from theories such as those of Hull (**8**) and Guthrie (**6**) which posit a conditioned antagonistic response as a partial or complete explanation for extinction. Such a response, if conditioned to the ready signal, would be expected to interfere (since it is antagonistic) with later performance in the presence of the CS.

A second possible negative effect of the ready signal is that, coming as it does just before the CS, it may function as an external inhibitor (**7**, p. 110). Although such an effect is transitory (**2**, p. 389), it might produce an artifactually low level of performance, at least on the early trials The fact that the no-ready-signal groups were superior to the ready-signal groups on Trials 1–20 supports such an interpretation. One important implication of this finding concerns the form of the acquisition function for eyelid conditioning. Spence (**16**), for example, has cited as supporting his theory the fact that such curves sometimes display an initial phase of positive acceleration. The data most favorable to this view, however, all came from experiments using a ready signal. Whether the same result will occur in studies using a less complicated procedure remains to be seen.

Summary

One hundred and twenty *Ss* participated in an eyelid-conditioning study where, after 20 ordinary conditioning trials on which the CS and UCS were paired, the experimental groups received a series of 20 trials on which the UCS was presented alone, and the control groups were given 20 more conditioning trials. The effect of this differential treatment was determined on a final series of 20 conditioning trials. The experiment was performed in two replications.

For half of the *Ss* in the experimental and control groups of each replication, each conditioning and UCS trial was preceded by a ready signal; for the other half, trials came without warning. Results, in agreement with those obtained in an earlier study, showed that, in terms of both performance on the final 20 conditioning trials and gain in performance during the middle block of 20 trials, the experimental and control groups with no-ready signal were comparable. The ready-signal experimental groups, however, were significantly inferior in performance to the ready-signal control groups. This latter result is like one recently reported by Goodrich, Ross, and Wagner, and identifies their use of a ready signal as the source of their failure to obtain the earlier results.

References

1. Boneau, C. A. The effect of changing the interstimulus interval on the conditioned eyelid response. Unpublished doctoral dissertation, Duke Univer., 1957.
2. Evans, C. A. L. *Recent advances in physiology.* London: Churchill, 1939.
3. Goodrich, K. P., Ross, L. E., & Wagner, A. R. Performance in eyelid conditioning following interpolated presentations of the UCS. *J. exp. Psychol.*, 1957, **53**, 214–217.
4. Grant, D. A., & Adams, J. K. 'Alpha' conditioning in the eyelid. *J. exp. Psychol.*, 1944, **34**, 136–142.
5. Grant, D. A., & Norris, E. B. Eyelid conditioning as influenced by the presence of sensitized Beta-responses. *J. exp. Psychol.*, 1947, **37**, 423–433.
6. Guthrie, E. R. *The psychology of learning.* (Rev. ed.) New York: Harper, 1952.
7. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: Appleton-Century-Crofts, 1940.
8. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943
9 Kimble, G. A., & Dufort R. H. The associative factor in eyelid conditioning. *J. exp. Psychol.*, 1956, **52**, 386–391.
10. Kimble, G. A., Mann, L. I., & Dufort, R. H. Classical and instrumental eyelid conditioning. *J. exp. Psychol.*, 1955, **49**, 407–417.
11. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
12. McAllister, W. R. The effect on eyelid conditioning of shifting the CS-UCS interval. *J. exp. Psychol.*, 1953, **45**, 423–428.
13. Passey, G. E. The influence of intensity of unconditioned stimulus upon acquisi-

tion of a conditioned response. *J. exp. Psychol.*, 1948, **38,** 420–428.
14. Pavlov, I. P. *Conditioned reflexes.* (G. V. Anrep, translator.) London: Oxford Univer. Press, 1927.
15. Spence, K. W. Learning and performance in eyelid conditioning as a function of intensity of the UCS. *J. exp. Psychol.*, 1953, **45,** 57–63.
16. Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.
17. Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45,** 116–119.
18. Underwood, B. J. *Psychological research.* New York: Appleton-Century, 1957.
19. Voeks, V. W. Acquisition of S-R connections: A test of Hull's and Guthrie's theories. *J. exp. Psychol.*, 1954, **47,** 137–147.

(Received October 21, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

DIFFERENTIAL EFFECTS OF "DISPLAY LAGS" AND "CONTROL LAGS" ON THE PERFORMANCE OF MANUAL TRACKING SYSTEMS [1]

W. D. GARVEY, J. S. SWEENEY, AND H. P. BIRMINGHAM

U. S. Naval Research Laboratory

In control systems, the arrangement of elements has no effect upon performance provided that the elements are linear and noise-free. For example, if three linear components such as an integrator, a lag network, and a lead network are to be cascaded, it can be shown mathematically that the transfer function of the combined unit will be independent of the particular order in which the components are arranged. On the other hand, if one of the elements of a system is either nonlinear, or noisy, or both (and the human, when operating in such a control system, may be so considered) the performance of the system will indeed be influenced by the order of the components through which the signals flow.

Although several studies (**1, 2, 3**) of the effects of low-pass filtering on compensatory target tracking have been reported, in none has the location of the filter, relative to the operator, been systematically varied. Yet, considerations of control theory would suggest that if a human and an electrical filter are to be cascaded in a system, the question of whether the human follows or precedes the filter is crucial in determining the effect of the filter upon system performance.

The extent to which the output of a tracking system follows the input is considered the ultimate measure of its performance. This measure of the difference between the input and output is called *system error* and it is the quantity which the system designer seeks to minimize.

It is the purpose of this study to test the effect which the location of a low-pass filter, relative to the subject, will have upon system performance. Figure 1 is a block diagram of the apparatus used. Merely by setting appropriate switches, a two-section, low-pass filter of adjustable constant could be inserted before or after the operator. The filter located ahead of the operator has come to be called a *display filter*, and one located after the operator is called a *control filter*.

METHOD

The *S*s performed a compensatory tracking task using a 5-in. CRT display and a position-control joy stick; i.e., a 1° displacement of the stick produced ¼-in. displacement of the target on the display. The target-course input consisted of a complex of three sine waves 3, 5, and 11 cycles per minute; amplitude of each was inversely proportional to its frequency. Four conditions of lag or low-pass filtering with time constants of .10, .24, .45, and .74 sec. (inserted either as display or control lags), and one condition of no lag were employed. The lags were created by using conventional two-section low-pass filters. The time constants of these filters were, in all cases, such as to provide negligible attenuation at the course frequencies. Eight naval enlisted men were each given one 1-min. trial on each condition per session in a counterbalanced order. The order was permuted from session to session through 10 sessions. The data for the last two sessions were used in the analysis of the results. Average system error was obtained for the last 50 sec. of the 60-sec. trials through the use of suitable full-wave rectification and a linear integrator.

[1] The authors are indebted to F. V. Taylor for his assistance in the interpretation of the results and the preparation of the manuscript.

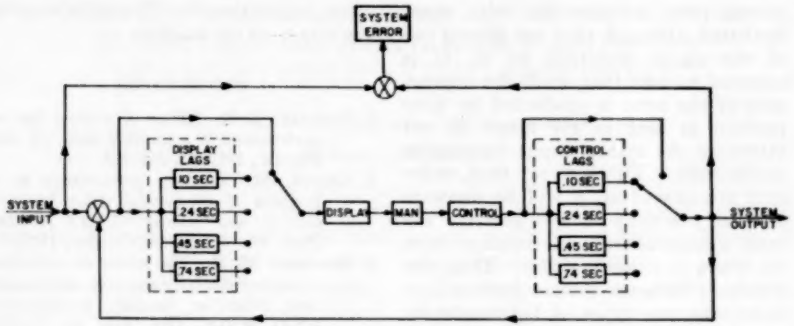

FIG. 1. Simplified block diagram of manual apparatus.

RESULTS AND DISCUSSION

The results are shown in Fig. 2. An analysis of variance did not show a significant effect from increasing the time constant of the control lag ($P > .05$); however, the effect of increasing the time constant of the display lag was found to be highly significant ($P < .001$). Using matched t tests the differences in performance between display lag conditions and control lag conditions for each time constant were compared; all differences were significant ($P < .01$) except the comparison for the .10-sec. lag ($P > .05$).

It is clear from the results that display lags produce a greater deterioration in system performance than do control lags. The functions shown in Fig. 2 would be altered under different conditions; e.g., with different course frequencies or stick sensitivities, control lags would probably be shown to affect system performance, and sometimes for the better. Thus, with greatly increased control gain and a long time constant low-pass filter, a situation similar to that of velocity control would be created with a resulting improvement in system performance for lower course frequencies (3). But the results as they are demonstrate clearly that the position of the filter in a human-operated tracking system is an important determinant of system performance.

The critical factor in this differential effect is the presence of a nonlinear component in the system emitting noise frequencies. In the present experiment this element was S; when a low-pass control filter is employed, noise is attenuated before the control output is added to the system input; thus the noise frequencies are diminished in the error, both as it is recorded and displayed to the tracker. Where a low-pass filter is inserted prior to the display, the

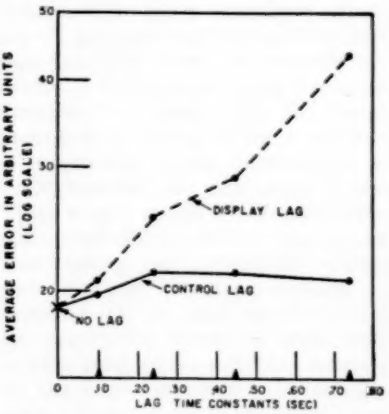

FIG. 2. Average error as a function of lag time constants (sec.). The parameters are the two lag conditions, display and control lags.

system error contains the noise terms unabated, although they are filtered out of the signal displayed to *S*. It is essential to note that, to *S*, the appearance of the error is unaffected by filter position so long as the filters do not attenuate the system input frequencies significantly. This is to say that, under such circumstances, *S* will be unaware of filter position; he will perceive the same display-control relationships with the filters in either position. Thus, the difference between system performance under the two types of lag conditions must be attributed to the different ways in which the machine components process *S*'s noise output relative to system error, rather than on *S*'s skill in keeping the target on the hairline.

References

1. Conklin, J. R. Effect of control lag on performance in a tracking task. *J. exp. Psychol.*, 1957, **53**, 261–268.
2. Levine, M. Tracking performance as a function of exponential delay between control and display. *USAF WADC Tech. Rep. No. 53-236*, October, 1953.
3. Rockway, M. R. The effect of variations in control-display ratio and exponential time delay on tracking performance. *USAF WADC Tech. Rep. No. 54-618*, December, 1954.

(Received July 17, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

INFLUENCE OF INTERTRIAL INTERVAL DURING EXTINCTION ON SPONTANEOUS RECOVERY OF CONDITIONED EYELID RESPONSES[1]

M. GORDON HOWAT AND DAVID A. GRANT

University of Wisconsin

The primary purpose of the present experiments was to investigate the effect of intertrial interval during extinction upon the subsequent spontaneous recovery of the conditioned eyelid response. Secondary observations were made on the influence of intertrial interval on the extinction process following spaced acquisition.

Greater spontaneous recovery has been expected following massed extinction as opposed to distributed extinction (**11,** p. 450), and both Hullian theory (**9,** pp. 75–77) and Estes probability theory of regression (**3, 4**) have predicted this result. According to Hull, massing the extinction trials when extinction is carried out to a criterion emphasizes the role of I_R in extinction, and highly distributed extinction trials emphasize the role of ${}_sI_R$ in the extinction process. Since I_R dissipates over time and ${}_sI_R$ does not, spontaneous recovery should be greater following massed extinction. The Estes theory would hold that spacing the extinction trials would result in the extinction of a greater proportion of the stimulus elements in the superpopulation and massed extinction would extinguish a less proportion of the elements in the superpopulation. Then, during the spontaneous recovery period following massed extinction trials, more conditioned stimulus elements could diffuse into the available subpopulation from the superpopulation which would result in a higher degree of spontaneous recovery.

The only previous study of the role of distribution of extinction trials on spontaneous recovery was that of Lewis (**10**) who studied the recovery of running time of rats in a 4-ft. runway. He found no significant effect of the distribution of extinction trials, although there was significantly greater spontaneous recovery following massed acquisition trials.

It is clear that further work is needed on this problem. The use of the human eyelid CR permits investigation of spontaneous recovery intervals longer than the 60 min. used by Lewis as the spontaneous recovery of the human eyelid conditioning would presumably be less influenced by the maintenance schedule of animals used in instrumental learning. To study spontaneous recovery in eyelid conditioning it seemed desirable to depart from the usual procedure of giving a fixed number of acquisition trials followed by a fixed number of extinction trials. This would result in considerable individual differences in amount of conditioning and amount of extinction. It was, therefore, decided to reinforce all *Ss* until they reached an acquisition criterion and then to extinguish all *Ss* until they reached an extinction criterion so that constant amounts of acquisition and extinction could be guaranteed before spontaneous recovery was sought.

[1] This research was supported in part by the Research Committee of the Graduate School of the University of Wisconsin with funds from the Wisconsin Alumni Research Foundation.

Procedure

Apparatus.—The basic apparatus employed was the same as that used by Grant, Schipper, and Ross (**8**). The *S* faced a circular milk-glass window 60 cm. directly in front of his eyes. He was instructed to watch the milk-glass window and not to interfere with his natural eyelid responses. The *S* wore a headgear on which was mounted the puff delivery mechanism and a microtorque potentiometer, the shaft of which was fastened to his right eyelid by means of a bit of piano wire and Scotch tape. The signal from the potentiometer was amplified and differentiated so as to give an ink recording of the eyelid movement and its first time derivative. The experiment was carried on under light-adapted conditions to minimize sensitization of the Beta-response (**7**). The CS consisted of an abrupt increase in the illumination of the milk-glass window from an initial level of 7 to a final level of 100 mL. The duration of the CS was 750 msec. The UCS was a corneal air puff administered to the left eye 500 msec. after the onset of the CS. The puff was generated by the fall of a 200-mm. column of mercury and was delivered through a glass tube of 3-mm. aperture. This intensity of the UCS was chosen because it had been shown to be nearly optimal in producing a high level of conditioning and resistance to extinction (**13**). The actual pressure at the eye is hard to specify but was much less than 200 mm. of mercury.

Procedure.—All *S*s received acquisition trials with a random intertrial interval averaging 35 sec. until they gave 8 CR's in 10 successive trials. After each *S* had reached the 8/10 criterion his CS was extinguished at one of two intertrial intervals to a criterion of five consecutive trials on which no CR occurred. He was then dismissed with instructions to return to the laboratory at a specified time. Upon his return *S* was tested for spontaneous recovery. The spontaneous recovery test consisted in giving extinction trials until the same criterion of no response on five consecutive trials was reached. All of the spontaneous recovery test trials were given with a random intertrial interval averaging 35 sec.

Experimental design.—A 2×2 factorial design was employed in which the intertrial interval during extinction was orthogonal to the length of the spontaneous recovery interval. The *S*s in the massed groups were given extinction trials with a random intertrial interval averaging 7.5 sec. The spaced groups were given extinction trials with a random intertrial interval averaging 35 sec. Half of the *S*s were tested for spontaneous recovery about 24 hr. after the end of the initial extinction period; the other half were tested 20 min. after the end of the initial extinction. The 24-hr. recovery data were obtained about six months prior to the 20-min. data. Owing to unforeseen results of the former experiment, the 20-min. recovery experiment was carried out. The homogeneity of acquisition and extinction results and error variance in the two experiments justified pooling all results.

Subjects.—The *S*s were 92 men and women students from introductory classes in psychology at the University of Wisconsin. They were randomly assigned to massed and spaced extinction groups with the restriction that the mean number of CR's to the criterion of acquisition be nearly equal for the two groups. There were 23 *S*s in each of the four experimental groups. Matching was done independently for the two groups given the 24-hr. recovery interval and the two groups given the 20-min. recovery interval. For 24-hr. recovery there were 12 men and 11 women in the spaced extinction group; 8 men and 15 women in the massed group. Both 20-min. recovery groups consisted of 9 men and 14 women. Procedural errors, failure of *S* to return for spontaneous recovery tests, or failure to meet acquisition or extinction criteria resulted in the loss of a few *S*s. No *S*s were discarded if they returned for the spontaneous recovery test.

TABLE 1

Mean Number of CR's and Mean Number of Trials to Criterion on Acquisition, Extinction, and Spontaneous Recovery Tests

Extinction	Acquisition		Extinction		Spontaneous Recovery	
	CR's	Trials	CR's	Trials	CR's	Trials
20-min. Recovery						
Spaced	11.20	19.7	8.11	17.4	4.72	12.78
Massed	11.05	20.6	12.76	24.74	4.35	12.35
24-hr. Recovery						
Spaced	12.09	22.70	6.52	15.78	5.35	13.91
Massed	12.13	24.70	12.35	25.09	4.57	12.65

Results

The results are reported in terms of two dependent variables: (*a*) the number of CR's to the criterion and

(*b*) the number of trials to the criterion. Occurrence of CR's during the second extinction session was interpreted as spontaneous recovery.

The essential data of the experiment are shown in Table 1. The mean number of CR's and the mean number of trials to the pertinent criterion are given for acquisition, extinction, and the spontaneous recovery test following the rest or recovery intervals of 20 min. and 24 hr. Examination of Table 1 reveals that the matching on the basis of CR's during acquisition was very close for both pairs of groups. The matching on the basis of average number of trials to criterion was also quite good, although no attempt was made to match on this criterion.

From Table 1 it is seen that during extinction the massed group gave 58% more CR's than the spaced group. The effect of intertrial interval during extinction was thus shown to have an immediately apparent effect which was in the opposite direction from that usually predicted.[2] Both the 24-hr. and 20-min. experiments gave this result. The over-all mean amount of spontaneous recovery was practically the same for the pairs of massed extinction and spaced extinction groups. There was about 20% more spontaneous recovery following 24 hr. than following 20 min. of postextinction rest which indicated that spontaneous recovery followed the conventional course of other experiments (**1, 2, 5, 6, 12**).

Essentially the same results were found for the mean number of trials to the criteria. The massed and spaced extinction groups differed considerably in mean number of trials to the extinction criterion in that the massed group required 44% more trials than the spaced group. In spontaneous recovery the mean number of trials to the criterion was essentially the same for both the massed and spaced groups, and there was greater spontaneous recovery following the 24-hr. interval than that following the 20-min. interval period.

TABLE 2

F's from Analyses of Variance of Number of CR's During Spontaneous Recovery Test (Column A) and CR's on Initial Extinction Minus CR's on Recovery Test (Column B)

Source	*df*	*F* Tests	
		A. Number of CR's	B. Extinction CR's Minus Recovery CR's
Distribution of initial extinction	1	0.01	4.26*
Duration of recovery interval	1	6.94**	1.37
Distribution × recovery interval	1	0.01	0.43
Error variance	88	(0.51)	(136.06)

* $P = .05$.
** $P = .01$.

Analyses of variance of the spontaneous recovery scores are summarized in Table 2. The number of CR's to the criterion are analyzed in Column A. In Column A the spontaneous recovery scores have been subjected to the cube-root transformation in order to normalize the error distributions. The greater recovery time produced a significantly greater mean recovery score. Distribution of trials during extinction failed to produce a significant difference, and the interaction likewise failed to reach significance.

[2] Lewis (**10**) summarizes the conflicting evidence on this point. As our consulting editor points out, however, most of the classical aversive conditioning results, including one from the Wisconsin laboratory (**8**), show more rapid extinction with massed trials, whereas animal runway experiments turn out otherwise.

Since the amount of spontaneous recovery was approximately equal in the massed and spaced groups as shown in Table 1 an attempt was made to normalize the initial extinction scores by subtracting, for each *S*, the number of CR's to the spontaneous recovery criterion from the comparable score during the first extinction. This, in effect, tended to control individual differences in responding in approximately the same fashion that an analysis of covariance would control these individual differences. The resulting error distributions were approximately normal. The variable analyzed in Column B was thus the number of CR's on first extinction minus the number of CR's on the spontaneous recovery test. This variable was found to be significantly affected by the distribution of trials during extinction, although the *F* was significant at only the .05 level. Recovery time and interaction were not significant for either score nor should they have been as they were, in effect, dummy variables. Precisely the same results were obtained by analysis of the trials to criterion scores, and similar statistical results were obtained with more conventional transformations.

Discussion

Spontaneous recovery.—The mean amounts of spontaneous recovery in terms of CR's or number of trials to the criterion in the spontaneous recovery test were so similar for the massed and spaced extinction groups that no significant difference would be expected even with a very large number of *S*s per group. These results agree precisely with those reported by Lewis (**10**) in his runway experiment with rats. It should be pointed out that the experiment was, in a sense, biased against this result. Because the initial acquisition was by means of distributed trials and the spontaneous recovery test was by distributed trials, the greater spontaneous recovery would be expected following massed extinction. This result was predicted because using the same intertrial interval during extinction as had been used during acquisition ought to extinguish precisely the same responses that were acquired with distributed training period. Massed extinction trials would presumably extinguish responses generalized from those acquired during spaced acquisition, and therefore the spontaneous recovery, if tested with spaced trials, should be greater following massed initial extinction. No such difference was found. Since a substantial number of *S*s were run in each group in the present experiment the results add force to those of Lewis, and it may be said that at the present time no one has demonstrated a significant effect of distribution of extinction trials upon subsequent spontaneous recovery.

The fact that greater spontaneous recovery was found following 24 hr. than following 20 min. is in line with the results of other experiments summarized by Lewis (**10**), and this result can be predicted on the basis of all major learning theories.

Rate of extinction.—The massed extinction groups gave approximately twice as many CR's and required approximately twice as many trials to extinguish as did the spaced extinction groups. (This unforeseen finding in the 24-hr. recovery experiment made it advisable to do the second 20-min. recovery experiment before publishing the results.) It was the case, however, that on the first six trials of initial extinction the spaced groups had a slightly higher average frequency of CR's than did the massed groups. Beyond six trials the mean amount of responding on all trials was considerably greater for the massed groups. If the amount of responding to a criterion is equated with response probability, then Estes' statistical theory of distributional phenomena in learning (**3, 4**) describes the results remarkably well. Estes (**4**, p. 376) predicted that the curves of re-

sponse probability for massed and spaced extinction groups would be expected to cross with the spaced group starting higher but overtaking the massed group and running below it during the later extinction period. Precisely this was found in first six trials for the 20-min. recovery groups, but the 24-hr. groups were identical on the first extinction trial, with the massed group extinction curve above the spaced curve on the remaining trials.

Although the greater resistance to extinction during massed extinction trials found in the present study may in some way be connected with the use of trials to criteria (or the specific criteria used) as contrasted with the use of a fixed number of trials, the finding was completely unexpected as it was not in line with classical aversive results (**10**). It resembles findings with instrumental reward training (**14, 15**).

Summary

A total of 92 human *Ss* were conditioned with distributed trials to a criterion of 8 CR's in 10 successive trials. Half the *Ss* were then extinguished with massed trials to a criterion of no response on five successive trials, and half were extinguished to the same criterion with distributed trials. These two groups were further subdivided, and half of each group was tested for spontaneous recovery following 20 min. and the other half following 24 hr. Two dependent variables were used: (*a*) the number of CR's to a criterion of five consecutive trials with no CR during spontaneous recovery tests and (*b*) the number of trials to reach this same criterion on the spontaneous recovery test.

The results indicated clearly that with spaced extinction there were significantly fewer CR's or trials required to reach the criterion of extinction than was the case with massed extinction trials. The mean amount of spontaneous recovery, in terms of either total number of CR's or trials to attain the criterion, was found to be essentially identical for massed or spaced extinction groups at both the 20 min. and 24 hr. recovery intervals. No major learning theory seems to predict this result.

References

1. Ellson, D. G. Quantitative studies of the interaction of simple habits: I. Recovery from specific and generalized effects of extinction. *J. exp. Psychol.*, 1938, **23**, 339–358.
2. Ellson, D. G. Spontaneous recovery of the galvanic skin response as a function of the recovery interval. *J. exp. Psychol.*, 1939, **25**, 586–600.
3. Estes, W. K. Statistical theory of spontaneous recovery and regression. *Psychol. Rev.*, 1955, **62**, 145–154.
4. Estes, W. K. Statistical theory of distributional phenomena in learning. *Psychol. Rev.*, 1955, **62**, 369–377.
5. Graham, C. H., & Gagne, R. M. The acquisition, extinction, and spontaneous recovery of a conditioned operant response. *J. exp. Psychol.*, 1940, **26**, 251–280.
6. Grant, D. A., Hunter, H. G., & Patel, A. S. Spontaneous recovery of the conditioned eyelid response. *J. gen. Psychol.*, in press.
7. Grant, D. A., & Norris, E. B. Eyelid conditioning as influenced by the presence of sensitized Beta-responses. *J. exp. Psychol.*, 1947, **37**, 423–438.
8. Grant, D. A., Schipper, L. M., & Ross, B. M. Effect of the intertrial interval during acquisition on the extinction of the conditioned eyelid response following partial reinforcement. *J. exp. Psychol.*, 1952, **44**, 203–310.
9. Hull, C. L. *Essentials of behavior.* New Haven: Yale Univer. Press, 1952.
10. Lewis, D. J. Acquisition, extinction, and spontaneous recovery as a function of percentage of reinforcement and intertrial intervals. *J. exp. Psychol.*, 1956, **51**, 45–53.
11. Miller, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.
12. Miller, N. E., & Stevenson, S. S. Agitated behavior of rats during experimental extinction and a curve of spontaneous recovery. *J. comp. Psychol.*, 1936, **21**, 205–232.
13. Prokasy, W. F., Jr., Grant, D. A., & Myers, N. A. Eyelid conditioning as a function of unconditioned stimulus intensity and intertrial interval. *J. exp. Psychol.*, 1958, **55**, 242–246.
14. Sheffield, V. F. Resistance to extinction as a function of distribution of extinction trials. *J. exp. Psychol.*, 1950, **40**, 303–315.
15. Stanley, W. C. Extinction as a function of the spacing of extinction trials. *J. exp. Psychol.*, 1952, **43**, 249–260.

(Received August 4, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

INDEPENDENCE IN THE LEARNING OF TWO CONSECUTIVE RESPONSES PER TRIAL

ARTHUR L. BRODY

Lehigh University

In this study an attempt was made to see whether there was independence in the learning of two consecutive responses on each trial, when each response was reinforced partially. By independence we mean that the first response per trial has no effect on and is not affected by the second response per trial. Previous studies (e.g., **1**) have applied statistical learning theory to situations in which one response is made by *S* on each trial, and that response is reinforced partially. In the present experimental situation, following a signal at the start of each trial, *S* made two consecutive predictions as to which one of two reinforcing events would occur. Each of *S*'s predictions was followed by the presentation of one of the two reinforcing events. The particular reinforcing event which occurred was controlled by *E* and was not contingent upon *S*'s response.

In terms of Estes' statistical learning theory, independence in this situation means that the sample of stimulus elements which may become conditioned to the first response per trial does not overlap the sample of stimulus elements which may become conditioned to the second response per trial. In other words, the two samples are independent. Labeling one of the reinforcing events an E_1, independence of the two samples of stimulus elements implies that if an E_1 follows the first response on a trial, a proportion of stimulus elements from one sample will become conditioned to that response, with a resultant increase in the probability of an E_1 prediction as the first response. If an E_1 follows the second response on a trial, a proportion of stimulus elements from a second sample will become conditioned to that response, with a resultant increase in the probability of an E_1 prediction as the second response. It follows that the asymptotic level of E_1 predictions depends upon the proportion of trials on which an E_1 occurs.

By choosing appropriate values for the probability with which an E_1 follows *S*'s first response on a trial and the probability with which an E_1 follows *S*'s second response on a trial, a test of the hypothesis of independence can be made. In this study, two groups of *S*s were under the same partial reinforcement schedule with respect to the first response per trial. It would, therefore, be expected that both groups would approach the same asymptotic level of E_1 predictions for the first response. Furthermore, it would be expected that the asymptotic level would match the probability of reinforcement of an E_1 prediction. With respect to the second response per trial, each group was under a different partial reinforcement schedule. It would be expected that the asymptotic level of E_1 predictions for the second response per trial should be different for the two groups and should match the corresponding reinforcement probabilities.

METHOD

Subjects.—The *S*s were 32 college students, 16 in each of two groups, designated Group I and Group II.

Apparatus.—The *S* sat facing a 2 × 2-ft. panel on which were mounted one green signal light in the center and two white reinforcing lights on either side of and slightly below the signal light. Directly below each of the white reinforcing lights was a telegraph key. On *E*'s side of the panel were two lights, each connected in series with one of *S*'s keys, that informed *E* which key *S* had pressed, a timing device that controlled the onset and duration of the signal light and reinforcing lights, and a switch to determine which of the two reinforcing lights would come on.

Procedure.—The *S*s were instructed as to what they were to try to predict, when they should make their predictions, and how the predictions were to be indicated. Each trial started with the presentation of the signal light which came on for 2 sec. Following the onset of the signal, *S* predicted which one of the two reinforcing lights would come on. The *S* indicated his prediction by pressing one of the two telegraph keys—the key below the light which he thought would go on. Simultaneously with the signal going off, one of the reinforcing lights came on and remained on for 2 sec. After the reinforcing light went off there was a 2-sec. interval, during which time *S* made a second prediction. This second prediction was indicated in the same way as the first prediction—by pressing one of the two keys. Following *S*'s second prediction, one of the two reinforcing lights came on and remained on for 2 sec. This was followed by a 4-sec. interval, and then the next trial began.

Four practice trials were given, using the same order of presentation of the reinforcing lights for all *S*s. For these trials the sequences used were left-left, left-right, right-left, and right-right in that order, the first term of each sequence referring to the reinforcing light which followed *S*'s first response on a trial and the second term referring to the reinforcing light which followed *S*'s second response on a trial.

Following practice, a series of 100 trials was given. The dependent variable was *S*'s prediction that one reinforcing light, an E_1, would come on. For one-half of the *S*s in Group I and in Group II, the left reinforcing light was designated E_1, and for the remaining half of the *S*s in each group, the right reinforcing light was designated E_1. The independent variable was the probability with which an E_1 would follow *S*'s first prediction on any trial and the probability with which an E_1 would follow *S*'s second prediction on any trial.

For both groups the probability that an E_1 followed the first prediction of any trial was set equal to .8. The probability that an E_1 followed the second prediction of any trial was set equal to .3 for Group I and was set equal to .6 for Group II. For each *S* two different random and independent reinforcement schedules were drawn, corresponding to the probability values with which an E_1 followed *S*'s first response on any trial and *S*'s second response on any trial. There was, then, partial reinforcement of an E_1 prediction as the first response of a trial and as the second response of a trial.

Results and Discussion

The hypothesis of independence leads to the prediction that the asymptotic level of E_1 responses depends only upon the probability with which an E_1 occurs. Since an E_1 follows the first response per trial with a probability of .8, the asymptote should approach .8 for both groups. Similarly, the asymptote for the second response per trial should approach .3 for Group I and .6 for Group II.

Figure 1 shows the proportions of E_1 predictions for blocks of 20 trials. The first response per trial is designated R_1 and the second response per trial is designated R_2. Group I is shown on the left and Group II is shown on the right.

To see whether the two groups differed with respect to the first response per trial, a t was obtained for the difference between the two groups on the total number of E_1 predictions over the 100 trials. No significant difference was found ($t = 1.38$, $P > .1$). Since we are more specifically interested in the asymptotic level of learning, the last block of 20 trials was selected as an estimate of the asymptote. The t between the two groups on the total number of E_1 predictions over these 20 trials was not significant ($t = 1.46$, $P > .1$). It may be concluded that, with respect to the first response per

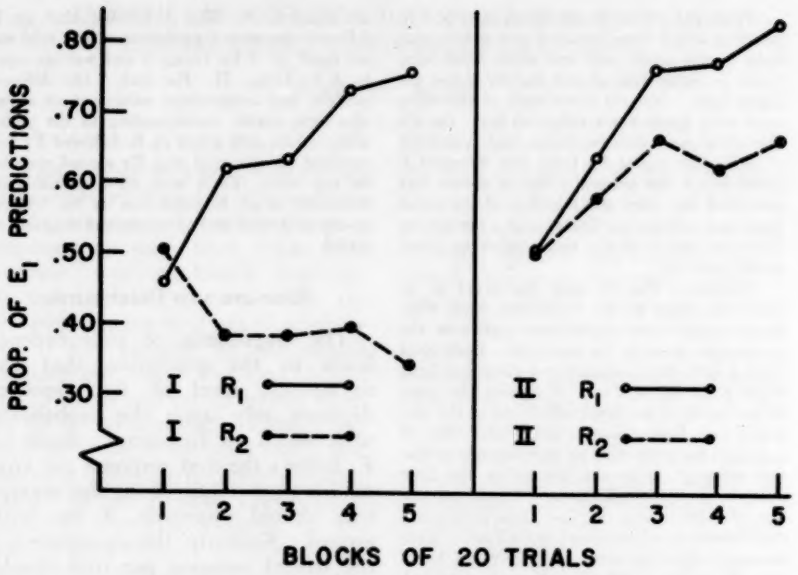

FIG. 1. Proportions of E_1 predictions for blocks of 20 trials.

trial, both groups were approaching the same asymptote. To see whether the observed asymptotes matched the probability of reinforcement of an E_1 prediction, t's were obtained for the differences between the observed and theoretical asymptotes for the first response per trial. For Group I, the observed asymptote was .76 and the theoretical asymptote was .80 ($t = .99$, $P > .3$). For Group II, the observed asymptote was .83 and the theoretical asymptote was .80 ($t = 1.12$, $P > .2$). It may be concluded that the asymptotic levels of E_1 predictions for the first response per trial matched the probability of reinforcement.

For the second response per trial, the t for the difference between the two groups on the total number of E_1 predictions over all trials was significant at the .001 level ($t = 5.70$). For the last block of 20 trials the t for the difference between the two groups on the total number of E_1 predictions was significant at the .001 level ($t = 4.17$). With respect to the second response per trial, the two groups were clearly approaching different asymptotes. The tests of the prediction that the asymptotes matched the probability of reinforcement of an E_1 prediction showed nonsignificant differences between the observed and theoretical asymptotes for the second response per trial. For Group I, the observed asymptote was .34 and the theoretical asymptote was .30 ($t = .92$, $P > .3$); for Group II, the observed asymptote was .67 and the theoretical asymptote was .60 ($t = 1.90$, $P > .05$). It may be concluded that the asymptotic levels of E_1 predictions for the second response per trial matched the probabilities of reinforcement.

It has been found that the asymptotic levels of learning for the first response per trial were the same for both groups and matched the probability of reinforcement, and that the asymptotic levels for the second response per trial differed for the two groups and matched the corresponding probabilities of reinforcement. It is therefore concluded that, in this situation, under the condition of partial reinforcement, the learning of two consecutive responses per trial proceeded independently, and that the asymptotic level of each response depended only upon the probability of reinforcement.

Two additional tests of the hypothesis of independence, though not included in the original design of the experiment, were carried out. To see whether the two groups differed with respect to the first response per trial, an orthogonal polynomial analysis-of-trends test (2) was run on the total number of E_1 predictions per block of 20 trials. The important comparison was the F for between-group means. No significant difference was found ($F = 1.92$, $P > .1$). This lends additional support to the hypothesis that the first response per trial was not significantly affected by the second response per trial.

If the two responses per trial were independent, the probability of a particular sequence of the two responses should depend only on the probability with which each response occurred. Since S predicted which one of two reinforcing events—an E_1 or an E_2—would occur, and since S made two such predictions per trial, there were four possible sequences of response. Either an E_1 or an E_2 prediction was made as the first response, followed by either an E_1 or an E_2 prediction as the second response. The frequency with which each sequence occurred over the last 20 trials was obtained for each of the two groups. These observed frequencies were compared with theoretical frequencies, based on the observed asymptotic values of E_1 and E_2 predictions. For each of the four sequences, the product of the two appropriate asymptotic values gave the probability of the sequence. These sequence probabilities, when multiplied by the total number of sequences on the last 20 trials, gave the theoretical frequency of each sequence. For each of the two groups, a χ^2 test was made between the observed and theoretical frequencies of the four sequences. For Group I, $\chi^2 = 3.84$ ($P > .2$) and for Group II, $\chi^2 = .74$ ($P > .8$). Since the differences between observed and theoretical frequencies of the four sequences were not significant, there is additional evidence that the first response of a trial had no significant effect on the second response of a trial. A similar comparison was made for the four possible sequences of S's second response on a trial followed by S's first response on the next trial. Either an E_1 or an E_2 prediction was made as the second response on a trial, followed by either an E_1 or an E_2 prediction as the first response on the next trial. For Group I, $\chi^2 = 1.85$ ($P > .5$) and for Group II, $\chi^2 = .04$ ($P > .99$). Since the differences between observed and theoretical frequencies of the four sequences were not significant, there is additional support for the hypothesis that the second response of a trial had no significant effect on the first response of the next trial.

Summary

In a two-choice situation *S*s made two consecutive predictions per trial, with each prediction followed by one of two reinforcing events. Two groups of *S*s were run under partial reinforcement schedules in which the reinforcing event was not contingent upon *S*'s prediction. The probability with which one of the reinforcing events, an E_1, followed *S*'s first prediction on any trial was set equal to .8 for both groups. The probability that an E_1 followed *S*'s second prediction on any trial was set equal to .3 for Group I, and equal to .6 for Group II. Both groups approached the same asymptotic level of E_1 predictions for the first

response per trial, but the two groups differed significantly with respect to the asymptotic level of E_1 predictions for the second response per trial. In all cases the asymptotic levels matched the corresponding probabilities of reinforcement. These results were interpreted as confirming the hypothesis that, in this situation, there was independence in the learning of two consecutive responses per trial, and that the asymptotic level of learning depended only upon the probability of reinforcement.

References

1. Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.
2. Grant, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.

(Received August 12, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

DISCRIMINATION LEARNING IN A VERBAL CONDITIONING SITUATION [1]

JULIET POPPER AND RICHARD C. ATKINSON
Stanford University *University of California, Los Angeles*

This study deals with an analysis of discrimination learning in a modified verbal conditioning situation (**1, 7**). Two stimuli, designated T_1 and T_2, are employed, and two responses, A_1 and A_2, are available to *S*. Each trial begins with the presentation of either T_1 or T_2. The probability of a T_1 trial is β and the probability of a T_2 trial is $1 - \beta$. Following the T_1 stimulus, an A_1 response is correct with probability π_1, and an A_2 response is correct with probability $1 - \pi_1$. Following the T_2 stimulus, an A_1 response is correct with probability π_2, and an A_2 response is correct with probability $1 - \pi_2$. The *S* is instructed to maximize the number of trials on which his response is correct.

Recently Burke and Estes have presented a component model for stimulus variables in discrimination learning (**2**). The theory is a direct extension of their original acquisition model (**4**). The only modification involves the definition of a set of stimulus elements to correspond to each of the stimulus situations to be discriminated (**5, 6**). Their theory is quantitative in a precise sense in that prior to experimentation exact predictions of behavior may be made for a wide range of parameter values.

One of the Burke and Estes discrimination problems deals with probabilistic reinforcement schedules which are contingent on the stimuli to be discriminated. The aim of the present study is to evaluate experimentally the adequacy of some of their predictions. In particular, we will be concerned with how adequately the theory accounts for asymptotic response behavior over a wide range of parameter values of π_2 given a fixed value of π_1.

The principal assumption in the Burke and Estes theory is that the stimuli T_1 and T_2 can be conceptually represented as two sets of stimulus elements which are designated S_1 and S_2, respectively. On T_1 trials *S*'s response is determined by a sample of stimulus elements from S_1; on T_2 trials, by a sample of stimulus elements from S_2. In addition, a set S_c is designated which represents those stimulus elements common to both sets S_1 and S_2 ($S_c = S_1 \cap S_2$), i.e., stimulus events common to the presentation of either T_1 or T_2. One can think of the expected size of the sample from the S_c set as providing an *index of similarity* between T_1 and T_2. The larger the relative size of this sample with respect to the expected size of the samples from S_1 and S_2 the greater the similarity between the stimuli (**3**).

Given the above stimulus representation and rules for conditioning (**2, 4**) one can derive an expression for the probability of an A_1 or A_2 response on Trial n. For our purposes we will be concerned with the long-

[1] This research was conducted at the Applied Mathematics and Statistics Laboratory, Stanford University, and was jointly supported by the Behavioral Sciences Division of the Ford Foundation and by the Group Psychology Branch of the Office of Naval Research. The authors are indebted to P. Suppes for several stimulating discussions of the ideas on which this experiment is based.

run probability of response behavior and, consequently, will present only asymptotic predictions; namely,

$$p_\infty(A_1|T_1) = \pi_1(1-\omega_1) + \pi_e\omega_1 \quad [1]$$

$$p_\infty(A_1|T_2) = \pi_2(1-\omega_2) + \pi_e\omega_2, \quad [2]$$

where $p_\infty(A_1|T_i)$ is the expected conditional asymptotic probability of an A_1 response given a T_i ($i = 1$ or 2) type trial, and $\pi_e = \beta\pi_1 + (1-\beta)\pi_2$.[2]

The quantity ω_i is a theoretical parameter and essentially represents a ratio of the number of elements in S_e to the number of elements in S_i.[3] Consequently $0 \leq \omega_i \leq 1$ and is independent of experimental parameters π_1, π_2, and β. In our situation T_1 and T_2 are symmetric, and it is natural to assume $\omega_1 = \omega_2 = \omega$. The closer ω is to unity, the greater the similarity between T_1 and T_2; the closer ω is to zero, the greater the dissimilarity between T_1 and T_2.

Method

Subjects.—The *Ss* were 205 undergraduates obtained from introductory courses in psychology. They were run in 20 subgroups, with from 8 to 14 *Ss* per subgroup.

Procedure.—Each *S* received a sheet of paper on which numbers from 1 to 320 were printed. Following each number were the letters A and B. The *Ss* were also provided with two blank sheets of heavy white paper. The instructions began as follows:

"I want to see how well you can do in a rather unusual problem situation. At the beginning of each trial, I will announce the number of the trial, and then I will say one of these two nonsense words: MEF or ZIL. About 4 sec. later I will say either A or B. Immediately after I have said ZIL or MEF on each trial, you are to guess whether I will say A or B on that trial by circling either A or B in the appropriate place on your answer sheet. If you expect an A, circle A; if you expect a B, circle B. Guess on every trial even if you are very unsure—your guesses or hunches may turn out to be right, and it is important to have a complete record of your learning. Try to improve as you go along and make as many correct choices as possible.

"I want you to make each choice without seeing any of your previous choices. Therefore, please take the small strip of heavy white paper and slide it down the answer sheet, covering each choice as soon as you have made it. When you have completed a column, use the big piece of paper to cover that whole column, and again use the small strip to slide down the next column." The remainder of the instructions involved repetition of the main points.

The *E* stood at the back of the room where it was possible to watch *Ss* to be sure they were following instructions. The number of each trial was announced; 1 sec. later *E* said either MEF or ZIL, 4 sec. after that either A or B, and 2 sec. later the number of the next trial was announced.

After Trial 4 *E* interrupted the sequence to ask whether *Ss* had any questions. Questions were answered by rephrasing appropriate parts of the instructions. The remaining 316 trials were run without interruption. The complete procedure took 45–50 min. for each subgroup.

MEF and ZIL were selected from the list of nonsense syllables with zero association value given by Glaze (**8**). They met the criteria of being clearly distinguishable, having an unambiguous correct pronunciation, and sounding unlike either A or B.

Design.—There were five experimental conditions. For all experimental groups π_1 was .85 and β was .50. The groups differed with respect to the π_2 parameter; the values of π_2 were .85 (Group I), .70 (Group II), .50 (Group III), .30 (Group IV), and .15 (Group V). Within each of the five experimental groups there were four subgroups distinguished as follows: 1. T_1 = MEF(T_2 = ZIL) and A_1 = A(A_2 = B); 2. T_1 = ZIL and A_1 = A; 3. T_1 = MEF and A_1 = B; 4. T_1 = ZIL and A_1 = B.

For each of the 20 subgroups a different random sequence of events was generated in accordance with the assigned values of π_1 and π_2, with the following restrictions imposed on the randomization: (*a*) each successive block of 40 trials included 20 T_1 trials and 20 T_2 trials; (*b*) of the 20 T_1 trials in a 40-trial block,

[2] These results, in slightly modified form, are presented in Equations 18 and 19 of the Burke-Estes article. In deriving these equations, Burke and Estes denote by S_i the set of stimuli sampled *only* on trials of type T_i. Our use of S_1 therefore corresponds to their use of $S_1 \mathbf{U} S_e$.

[3] To be exact, $\omega_i = \dfrac{N_e\theta_e}{N_e\theta_e + N_i\theta_i}$ for i = 1 or 2. N_e, N_1, and N_2 are the number of stimulus elements in the sets S_e, $(S_1 - S_e)$, and $(S_2 - S_e)$, respectively. θ_e, θ_1, and θ_2 are the sampling constants associated with these sets.

an A_1 response was correct on exactly $\pi_1 \cdot (20)$ of the trials and an A_2 was correct on the remaining $(1 - \pi_1) \cdot (20)$ trials, and (c) of the 20 T_2 trials in a 40-trial block, an A_1 response was correct on exactly $\pi_2 \cdot (20)$ of the trials and an A_2 was correct on $(1 - \pi_2) \cdot (20)$ trials.

Groups I, II, IV, and V consisted of 40 *S*s each; there were 45 *S*s in Group III.

Results and Discussion

Figure 1 presents the mean group response behavior over all trials of the experiment. In this figure the proportion of A_1 responses given a T_1 trial and the proportion of A_1 responses given a T_2 trial are plotted in successive blocks of 40 trials. In each block of 40 trials there are 20 T_1's and 20 T_2's; therefore the proportion computed for an individual *S* is based on 20 observations.

An inspection of this figure indicates that response curves are fairly stable over the last 120 trials, and it appears reasonable to assume that a constant level of responding has been reached. Consequently, proportions computed over the last 120 trials were used as estimates of $p_\infty(A_1|T_1)$ and $p_\infty(A_1|T_2)$.

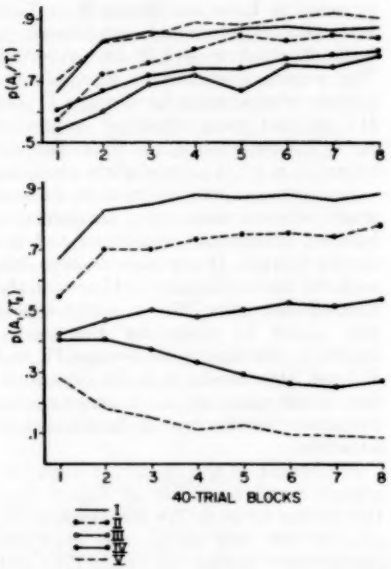

Fig. 1. The observed mean values of $p(A_1|T_1)$ and $p(A_1|T_2)$ in successive blocks of 40 trials for Groups I–V. Each point is based on 20 observations per *S*.

TABLE 1

Observed Values of $p_\infty(A_1|T_1)$ and $p_\infty(A_1|T_2)$ Computed over Last 120 Trials[a]

Group	$p_\infty(A_1\|T_1)$		$p_\infty(A_1\|T_2)$	
	Mean	*SD*	Mean	*SD*
I	.871	.098	.870	.092
II	.878	.099	.766	.105
III	.757	.119	.515	.157
IV	.785	.139	.262	.172
V	.910	.085	.088	.087

[a] Each entry based on 60 observations per *S*.

A simple analysis of variance for each experimental group tested for differences between the four subgroups with respect to the total number of A_1 responses in the last 120 trials. None of the five analyses was significant at the .05 level, and for subsequent analyses the subgroup distinctions within an experimental group were not considered.

Table 1 presents the observed means and *SD*s of proportions for the last 120 trials. Entries for Groups I, II, IV, and V are based on $N = 40$; entries for Group III are based on $N = 45$.

Figure 2 presents a schematic representation of the theoretical predictions as well as the observed asymptotic values given in Table 1. By substituting appropriate parameter values in Equations 1 and 2 the following results are established: (a) If $\omega = 0$, then (for the group designation on the abscissa) $p_\infty(A_1|T_1)$ falls on the straight line ab and $p_\infty(A_1|T_2)$ falls on the straight line ad. (b) If $\omega = 1$, then both $p_\infty(A_1|T_1)$ and $p_\infty(A_1|T_2)$ fall on the straight line ac. (c) If $0 < \omega < 1$, then (i)

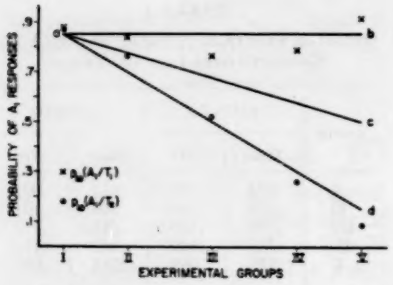

FIG. 2. A schematic representation of theoretical and observed asymptotic predictions.

$p_\infty(A_1|T_1)$ falls on a straight line with origin at point *a* and bounded between lines *ab* and *ac* and (*ii*) $p_\infty(A_1|T_2)$ falls on a straight line with origin at point *a* and bounded between lines *ac* and *ad*. Further, the amount of displacement of $p_\infty(A_1|T_1)$ from the line *ab* is the same as the displacement of $p_\infty(A_1|T_2)$ from the line *ad* and is a function of the value of ω.

Inspection of Fig. 2 suggests certain analyses of the data. In particular, the expected values of $p_\infty(A_1|T_1)$ for the various groups should lie on or between the lines *ab* and *ac*. However, the observed values for both Groups I and V are above the *ab* line. To establish whether these observed points are significantly above the maximum value of .85, *t* tests were run employing the observed *SD*s of the mean as the error term. The obtained values of *t* were 1.36 and 4.46 ($P < .01$, 39 *df*) for Groups I and V, respectively. Thus, for Group V the observed value of $p_\infty(A_1|T_1)$ was significantly greater than the maximum value predicted by the theory.

Similarly the expected values of $p_\infty(A_1|T_2)$ for the various groups should lie on or between the lines *ac* and *ad*. Yet for Group I the observed value is above .85 and for Groups IV and V the observed values are below .30 and .15, respectively. To establish whether the observed results were significantly different from these values, *t* tests were again run. The obtained values of *t* were 1.37, 1.40, and 4.58 ($P < .01$, 39 *df*) for Groups I, IV, and V, respectively. Thus, for Group V the observed value of $p_\infty(A_1|T_1)$ is significantly smaller than the minimum value predicted by the theory.

A more stringent requirement of the theory is that (*a*) the expected values of $p_\infty(A_1|T_1)$ fall on a straight line bounded between the lines *ab* and *ac* and (*b*) the expected values of $p_\infty(A_1|T_2)$ fall on a straight line bounded between the lines *ac* and *ad*. Inspection of Fig. 2 clearly indicates, at least for $p_\infty(A_1|T_1)$, that this is not the case. The observed values of $p_\infty(A_1|T_1)$ decrease from Groups I to III but, instead of continuing in this trend, show a marked increase for Groups IV and V. Thus we find a convex rather than a linear relationship between $p_\infty(A_1|T_1)$ and the groups when ordered from I to V.

An experiment employing a procedure similar to the one used in this study was reported by Estes and Burke (**5**). They had two experimental conditions, in both of which $\pi_1 = 1.00$ and $\pi_2 = .50$. The results qualitatively resemble the present observations for Groups II and III in that their observed values of $p_\infty(A_1|T_1)$ are below π_1 and their observed values of $p_\infty(A_1|T_2)$ are slightly above π_2.

In conclusion, the results of the present study indicate substantial disagreement between theoretical predictions and observed values. If one were dealing only with the data of Groups I–III or with the Estes-Burke data (**5**), a fairly strong case could be made for the model. However, the results on Groups IV and V leave little doubt that the formalization is not adequate in its present form to account for this type of discrimination situation.

In support of the theory, it might be argued from inspection of Fig. 1 that the curves of $p(A_1|T_1)$ for Groups III and IV are still rising, and that the asymptotic values of $p(A_1|T_1)$ and $p(A_1|T_2)$ are in reality π_1 and π_2, respectively, for all groups. As noted above, the theory would predict these asymptotic values when $\omega = 0$. However,

if $\omega = 0$, it necessarily follows from the theoretical formulation that the curves of $p(A_1|T_1)$ for all five groups should be identical over trials, approaching π_1 at the same rate. An inspection of the upper panel of Fig. 1 clearly indicates that this is not the case.

If, in fact, $p(A_1|T_1)$ and $p(A_1|T_2)$ are asymptotically equal to π_1 and π_2, respectively, but are approaching the asymptotes at differential rates depending on both π_1 and π_2, one might conjecture that stimuli common to T_1 and T_2 trials (S_c) are initially influencing behavior but are becoming adapted over trials, i.e., losing their influence on response probability (**9**). The rate of adaptation could be assumed to be a function of the π values. This theoretical approach appears to be worth investigating.

It should be noted that, in one sense, this type of study cannot be viewed as a satisfactory test of the theory. It may be that the relationship of asymptotic response behavior to π_1 and π_2 obtained in the present experiment is an artifact of the massed trial procedure employed. In the derivation of Equations 1 and 2 it was assumed by Burke and Estes that either S_1 or S_2 was sampled *independently* on each trial. If, however, the trials occur in close temporal succession, then the stimulus complex affecting S on any trial may include traces of the stimulation associated with the responses and reinforcing events of one or more preceding trials. Therefore, a more acceptable test of the theory, in the form presented by Burke and Estes, would require a situation where experimental techniques are employed to reduce the carry-over of trace stimulation from one trial to the next (**10**).

Summary

The study deals with an analysis of discrimination learning. Two stimuli designated T_1 and T_2 are employed and two responses A_1 and A_2 are available. Each trial begins with the presentation of either T_1 or T_2, the probability of T_1 and of T_2 each being $\frac{1}{2}$. Following T_1, an A_1 response is correct with probability π_1 and an A_2 response is correct with probability $1 - \pi_1$. Following T_2, an A_1 response is correct with probability π_2 and an A_2 response with probability $1 - \pi_2$. The S is instructed to maximize the number of trials on which his response is correct.

Five groups were run. For all groups $\pi_1 = .85$. The groups differed with respect to the π_2 parameter which assumed the values .85, .70, .50, .30, and .15. Analysis of the data was in terms of a theoretical model for discrimination learning proposed by Burke and Estes (**2**). Discrepancies between predicted and observed outcomes were examined. However, it was pointed out that the massed trial procedure employed in the study did not provide an optimal test of the theory.

References

1. Atkinson, R. C. An analysis of the effect of nonreinforced trials in terms of statistical learning theory. *J. exp. Psychol.*, 1956, **52**, 28–32.
2. Burke, C. J., & Estes, W. K. A component model for stimulus variables in discrimination learning. *Psychometrika*, 1957, **22**, 133–145.
3. Bush, R. R., & Mosteller, F. A model for stimulus generalization and discrimination. *Psychol. Rev.*, 1951, **58**, 413–423.
4. Estes, W. K., & Burke, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–286.
5. Estes, W. K., & Burke, C. J. Application of a statistical model to simple discrimination learning in human subjects. *J. exp. Psychol.*, 1955, **50**, 81–88.
6. Estes, W. K., Burke, C. J., Atkinson, R. C., & Frankmann, J. P. Probabilistic discrimination learning. *J. exp. Psychol.*, 1957, **54**, 233–239.
7. Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.
8. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–267.
9. Restle, F. Theory of selective learning with probable reinforcements. *Psychol. Rev.*, 1957, **64**, 182–191.
10. Straughan, J. H. An application of statistical learning theory to an escape learning situation using human subjects. *J. exp. Psychol.*, 1956, **52**, 1–8.

(Received August 12, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

EFFECT OF THE SPATIAL RELATIONSHIP BETWEEN CUE, REWARD, AND RESPONSE IN SIMPLE DISCRIMINATION LEARNING[1,2]

J. V. MURPHY AND R. E. MILLER

School of Medicine, University of Pittsburgh

The spatial relationship between the cue and the rewarded response in discrimination learning has recently been studied by several investigators (**8, 11**). In such studies a vertical displacement of the cue from the rewarded response resulted in a highly significant retardation of discrimination learning. An additional study (**10**) investigated the possibility that a discrimination problem presented in a vertical plane might account for the sensitivity of the spatial variable. Discrimination learning was found to be unaltered by the plane of presentation of the discriminanda.

In the earlier studies by McClearn and Harlow (**8**) and by Murphy and Miller (**11**), the cue was vertically displaced from the reward and response—i.e., the reward and response were spatially contiguous. The present study was designed to investigate all possible combinations of a spatial separation between the cue, reward, and response in simple discrimination learning. A 7-in. separation was the unit of displacement investigated.

METHOD

Subjects.—The *S*s were five mature male rhesus monkeys. These animals had previously been used in a study of interanimal conditioning and dominance behavior (**12**). They were given extensive adaptation training in the test situation. In the preliminary training phase *S* learned to displace a response block to obtain a raisin reward located directly beneath it.

Apparatus.—The Wisconsin General Test Apparatus (WGTA) was employed (**3**, Fig. 1). This consisted essentially of a restraining cage with bars on one side through which *S* could reach and displace objects on a movable tray. A special form board was constructed (see Fig. 1) which contained two trays with one located 7 in. above the other.

One tray contained two foodwells located 3½ in. from each side of the tray and 14 in. apart center to center. Each foodwell was equipped with a sliding micarda cover. This slide was so constructed as to be flush with the top of the tray and was retracted by *E* to expose the foodwell when a correct response was made. Two inches behind each foodwell was located a magnetic cabinet latch. The second tray contained a single foodwell located in the center of the tray. This foodwell was likewise equipped with a sliding cover. Magnetic cabinet latches were mounted in the same relative positions as on the two-hole tray—i.e., on each side 14 in. center to center. The trays were interchangeable between the top or bottom shelves, and when inserted in the proper position the shelves were locked into position.

The discriminanda were 2 × 2 × ¾-in. wooden blocks, one painted white and the other black. A metal brace was attached to the back of each object. When in position, the brace was held magnetically by the cabinet latches thereby suspending the stimuli ¾ in. directly above the foodwells on the two-hole tray or in a comparable position on the one-hole tray. The stimulus objects were interchangeable in either right or left position on both trays. The 2 × 2 × ¾ in. response blocks were unpainted and could be placed under the stimulus blocks. The response blocks were free to move in any direction while the stimuli remained stationary.

The form board (Fig. 1) containing the trays was 16 in. high and completely enclosed except for the front side. The micarda foodwell covers projected through the back wall and could be easily manipulated by *E*. The form board was built on a "Lazy Susan" design which facilitated stimulus and baiting changes. The one-way vision screen in the WGTA was not

[1] This investigation was supported by a research grant (M-487C) from the National Institute of Mental Health of the National Institutes of Health, U. S. Public Health Service.

[2] The authors are indebted to I. Arthur Mirsky for suggestions and criticisms.

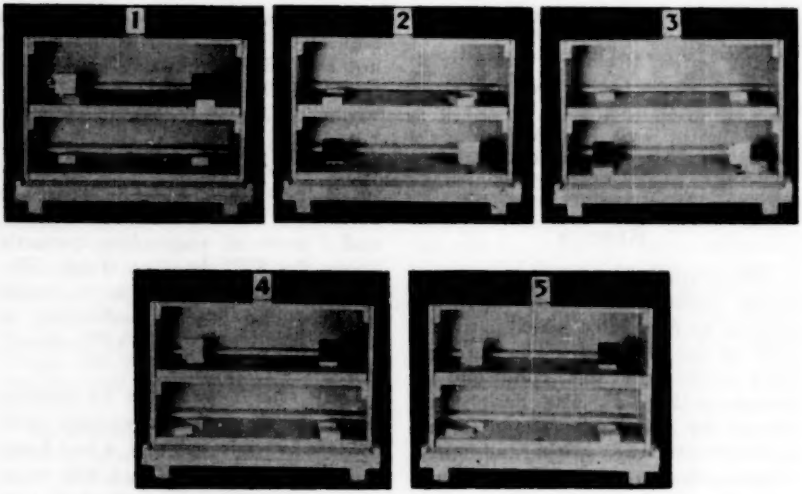

FIG. 1. The form board with the experimental conditions as they appear after a correct choice has been made. The open foodwell indicates the location of the reward. The numbers refer to the five experimental procedures described in the text.

employed since the subjects were completely removed from view by the enclosed form board. The *S*'s choice was observed through a mirror placed above the opaque screen. This mirror was inaccessible to *S*'s view. Any movement by *S* of a response block constituted a trial. A noncorrection procedure was employed throughout. In summary form, the modifications of the WGTA entailed the following: (*a*) two trays were employed, (*b*) the foodwells were covered, (*c*) the discriminanda remained stationary when a response occurred, and (*d*) Ss were observed through a mirror arrangement.

Procedure.—The problem for all variations of spatial separation was a simple black-white discrimination. White was the positive or rewarded stimulus throughout. To solve any arrangement of the problem it was merely necessary for *S* to move the response block on the same side as the white stimulus block. The position of the positive stimulus was randomly varied from right to left side. In the test situation, an opaque screen was lowered during the baiting. After the raisin reward had been placed in the foodwell, the foodwell cover was closed. Then the opaque screen was raised and the stimulus tray pushed just within reach of *S*. If *S* made a correct choice—i.e., displaced the response block located either directly below or 7 in. below the white stimulus object—the appropriate foodwell cover was immediately withdrawn by *E*. It may be noted that the movement of the foodwell cover never produced a discernible reaction on the part of *S* even during preliminary training.

Each *S* was given 400 trials on each of five tasks. Fifty trials were run daily with 25 trials in an A.M. testing session and 25 in the P.M. The conditions of spatial separation are shown in Fig. 1. Thus in Cond. I, the cue and response were spatially contiguous and the reward was 7 in. removed; in Cond. 2, the cue and reward were contiguous and the response was removed 7 in. In Cond. 3, the cue, reward, and response were all contiguous. Since this condition required that *S* attend to only one of the two trays, the control of running half the trials from the upper tray and the remaining half from the lower tray was initiated. Accordingly, in the A.M. the trials were conducted from the top shelf and those in the P.M. from the bottom shelf. In Cond. 4, the reward and response were contiguous with the cue spatially removed; in Cond. 5 the spatial factor was introduced between all three aspects. Since in Cond. 5 it was impractical to locate the reward 7 in. from both the cue and response at the same time, as this would have required a foodwell located between the two trays, the reward was located 7 in. from the response and 10 in. from the cue in the A.M. as illustrated in Fig. 1. In the P.M. the response block was moved to the

top tray with the stimulus block on the lower tray. That is, in the P.M. the reward was 7 in. away from the cue and 10 in. from the response.

The order of presentation for successive variations of the problem was determined by a randomly drawn latin square. Thus all *Ss* were ultimately tested on all five variations of the problem. The *Ss* were tested on successive days with no interval between tasks.

Results

The over-all analysis of variance of errors under the various conditions yielded an F of 12.61 which is significant at the .01 level. The performance under the five conditions is shown in Fig. 2. The design employed for the study was a Treatments × *Ss* design with randomized sequence and order effects (7). Such a design eliminates any systematic effects due either to order or sequence of treatments. Thus, in the over-all analysis these aspects are controlled; however these factors are confounded when the comparison of any two tasks is attempted. In order to determine which conditions of spatial contiguity were learned, the Alexander trend test (1) was employed. This test indicated a significant improvement with practice for *Ss* tested under Cond. 1, 2, and 3. It is clear that Cond. 1, 2, and 3 were learned to a high level, while Cond. 4 and 5 were not learned. Supporting evidence is found in the "direct" effects of the various treatment conditions. The "direct" effect refers to the effect of a treatment when it is administered first for a particular *S*. The three *Ss* performing first under Cond. 1, 2, and 3 were all responding correctly above the 80% level on Trials 376–400. In contrast, the two *Ss* under Cond. 4 and 5 were performing at chance levels (48% and 52% correct, respectively).

Only one *S* was able to perform above 80% correct responses over the last 25 trials on Cond. 4 and 5 and this occurred after *S* had 800 trials of previous experience on learnable arrangements of the problem.

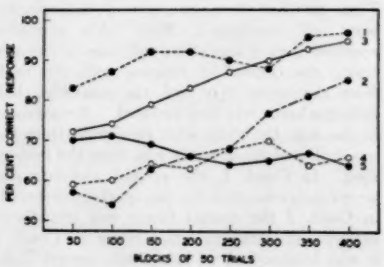

FIG. 2. The percentage of correct response for the various conditions: (1) cue, response . . . reward; (2) cue, reward . . . response; (3) cue, reward, response; (4) cue . . . reward, response; (5) cue . . . reward . . . response.

Discussion

The present study was designed to investigate the importance of the spatial relationships between the cue, the reward, and the site of the response in discrimination learning. Since previous work (11) had indicated that learning is minimal or nonexistent when the cue is spatially removed from the rewarded response (Cond. 4), the question arose as to whether it was the spatial separation of the cue from the reward (Cond. 1) or the spatial separation between cue and response (Cond. 2) which was the crucial factor in producing the earlier results. In the present study learning was observed to occur in both Cond. 1 and 2; consequently, it appears that neither factor alone disrupts learning but when both are spatially removed from the cue, learning is significantly retarded. It is evident that learning occurred in the three situations in which the cue was contiguous with either the response or the reward but did not occur in the two situations in which the cue was not spatially contiguous with the reward or response.

The significantly poorer performance in Cond. 4 and 5 as compared with Cond. 1, 2, and 3 is particularly striking in view of the transfer of training which might be expected in this type of experiment. Some of the *S*s had successfully solved one or more of the spatial arrangements of the discrimination problem and would, therefore, be expected to demonstrate some positive transfer effects. Despite these favorable conditions, performance on Cond. 4 and 5 showed no improvement over the course of 400 trials. A similar finding that spatial discontinuity disrupted a previously learned discrimination habit was reported in an earlier experiment (**11**).

The sequence of tasks for the one *S* which successfully solved all five variations of the problem was 2, 1, 4, 5, 3. It was particularly interesting to note that this *S* solved Cond. 4 and 5 by using a double response. That is, *S* would touch the positive stimulus and then move down 7 in. and move the response block at which time the reward was made accessible. This practice was particularly apparent under Cond. 5. After *S* was performing successfully with this technique for a series of trials the stimulus touching response was gradually reduced to a movement toward the stimulus without actual contact. Later in the series of trials *S* began to respond directly to the response block. At this point his performance deteriorated and the number of errors increased until he once more started utilizing the double response procedure. It is this phenomenon which accounts in large part for the unusual decrease in the percentage of correct responses with an increasing number of trials shown in Fig. 2 for Cond. 5.

Jarvik (**4**) recently reported an experiment in which chimpanzees had considerably more difficulty in forming a discrimination when the reward was located ¼ in. below the positive stimulus than when the reward was placed within the stimulus itself. The suggestion was made that the spatial displacement of the reward may be sufficient to impair performance when only a ¼ in. displacement occurs. In Cond. 1 of the present study the reward was spatially removed 7 in. from the cue and response. The high degree of learning obtained under this condition suggests that the differences observed by Jarvik were attributable to some other factor. In the Jarvik experiment the reward within the stimulus condition was obtained by placing the reward in the bottom of the stimulus object with Scotch tape placed across the opening. The *S* was required to remove the tape to procure the reward. It appears probable that the differences obtained by Jarvik were attributable to the additional manipulation of the stimulus when the reward was placed inside or to some other such phenomenon rather than the spatial factor.

Laboratory studies of discrimination learning provide an example of the almost completely unrecognized possibility of misinterpretation attributable to the spatial parameter in many areas of psychological investigation. The Wisconsin General Test Apparatus is employed extensively for discrimination learning with monkeys. In this situation the cue, reward, and response are spatially contiguous (Cond. 3). An apparatus employed extensively with rats is the Lashley Jumping Stand (**6**). In this case the cue and reward are contiguous with the response removed some 9 to 12 in. (Cond. 2). Although learning would be predicted in both instances on the basis of the present study, it appears clear that the two situations may differ considerably. It is noteworthy that the Yerkes Apparatus (**15**) which was particularly inefficient for discrimination in rats is analogous to our Cond. 5.

The possibility of misconceptions attributable to the spatial factor is apparent in several recent studies (**5, 13**) concerned with the investigation of Harlow's learning-set phenomenon in humans and rats. Since the learning-set phenomenon has been extensively studied

in the monkey, it appears that such studies are an attempt to provide comparative data at the higher and lower levels of the scale. Although these studies will not be discussed in detail, it is pertinent to note that neither employ a situation in which the cue, reward, and response are contiguous as in the original work by Harlow (2). On the other hand Warren and Baron (14), and Miles and Meyer (9) have provided excellent comparative data without introducing any change in spatial relationships from those used by Harlow (2).

It is also interesting to consider maze learning in terms of the spatial parameter. The maze situation is somewhat similar to Cond. 1 in which the cue and the response are spatially contiguous with the reward removed spatially. However, to be analogous it is required that *S* utilize cues which are much more contiguous with the response than has generally been recognized. If *S* were utilizing cues which are spatially removed from the choice-point response, the maze situation would be analogous to our Cond. 5 and essentially unlearnable. This suggests the somewhat unusual hypothesis that, other factors being equal, the suspension of prominent differential stimuli above the correct alternatives at the choice points will not facilitate acquisition of maze behavior. The implications of the spatial parameter in conditioning studies has previously been discussed in some detail (8) and is currently under investigation.

It is clear that considerable work remains to be completed in this area. An important problem is an extensive test to determine the conditions under which a significant difference occurs between the three learnable arrangements of the problem. The present study was designed to delineate those aspects of the spatial parameter which merit additional study with subhuman primates and those aspects which may merit further study as a variable for detecting developmental levels in children.

Summary

The spatial relationships between the cue, reward, and response in simple black-white discrimination were investigated. A 7-in. vertical distance was the unit employed. Cue, response, and reward were varied in all of the five possible combinations of contiguity-separation.

The five rhesus monkeys were tested on each condition in the Wisconsin Apparatus. A special form board enabled the independent manipulation of cue, response, and reward. Order of presentation was determined by a randomly drawn latin square. An analysis of variance of errors under the various arrangements of the problem was significant ($P = .01$). Learning occurred in the three situations in which the cue was contiguous with either the response or the reward but did not occur in the two situations in which the cue was not spatially contiguous with the reward or response.

References

1. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
2. Harlow, H. F The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51–65.
3. Harlow, H. F., & Bromer, J. A. A test apparatus for monkeys. *Psychol. Rec.*, 1938, **2**, 434–436.
4. Jarvik, M. E. Simple color discrimination in chimpanzees: Effect of varying contiguity between cue and incentive. *J. comp. physiol. Psychol.*, 1956, **49**, 492–495.
5. Koronakos, C., & Arnold, W. J. The formation of learning sets in rats. *J. comp. physiol. Psychol.*, 1957, **50**, 11–14.
6. Lashley, K. S. The mechanism of vision. I. A method for rapid analysis of pattern vision in the rat. *J. genet. Psychol.*, **37**, 453–460.
7. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.
8. McClearn, G. E., & Harlow, H. F. The effect of spatial contiguity on discrimination learning by rhesus monkeys. *J. comp. physiol. Psychol.*, 1954, **47**, 391–394.
9. Miles, R. C., & Meyer, D. R. Learning sets in marmosets. *J. comp. physiol. Psychol.*, 1956, **49**, 219–223.
10. Miller, R. E., & Murphy, J. V. Discrimination learning with vertical vs.

horizontal stimulus relationships. *J. comp. physiol. Psychol.*, 1956, **49**, 80–83.

11. MURPHY, J. V., & MILLER, R. E. The effect of spatial contiguity of cue and reward in the object-quality learning of rhesus monkeys. *J. comp. physiol. Psychol.*, 1955, **48**, 221–224.
12. MURPHY, J. V., MILLER, R. E., & MIRSKY, I. A. The manipulation of dominance in monkeys with conditioned fear. *J. abnorm. soc. Psychol.*, 1956, **53**, 244–248.
13. SHEPARD, W. O. Learning set in preschool children. *J. comp. physiol. Psychol.*, 1957, **50**, 15–17,
14. WARREN, J. M., & BARON, A. The formation of learning sets by cats. *J. comp. physiol. Psychol.*, 1956, **49**, 227–231.
15. YERKES, R. M., & WATSON, J. B. Methods of studying vision in animals. *Behav. Monogr.*, 1911, **1**, No. 2.

(Received August 19, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

DELAY OF REINFORCEMENT, RESPONSE PERSEVERATION, AND DISCRIMINATION REVERSAL [1]

BENJAMIN H. PUBOLS, JR. [2]

University of Wisconsin

The empirical relationship between delay of reinforcement and acquisition of instrumental responses in rats has been well established. In a variety of learning situations, the typical finding is that acquisition is a negatively accelerated, decreasing function of the length of delay in reinforcement (**5, 12, 13, 25**). However, there is another functional relationship involving delay of reinforcement that has until recently been relatively neglected in research. This is the relationship between delay of reinforcement and resistance to experimental extinction, or response perseveration, in the absence of reinforcement.

Since 1951, several studies have appeared employing one or both of two delay procedures, a constant delay over all acquisition trials for a given group of *Ss* (**4, 11**), or a variable delay for a given group (**1, 10, 11, 14, 17**). Results of these studies may be briefly summarized by stating that, in general, resistance to extinction is greater following constant or varied delay than following no delay. The only exception seems to be in the first of two studies reported by Logan, Beier, and Kincaid (**11**), who found insignificant differences in extinction of a running response in rats following 1-sec. constant delay, 5-sec. constant delay, and 1-sec. vs. 9-sec. variable delay.

Two of the studies using the variable delay procedure (**1, 17**) were designed and interpreted in relation to the Hull-Sheffield hypothesis regarding resistance to extinction following intermittent reinforcement (**18**). For example, Crum, Brown, and Bitterman (**1**) argue for the importance of events occurring on trials on which reinforcement is withheld, rather than on trials following them. Their argument is based on the fact that, in their study, aftereffects of reinforcement were present on all trials, and yet the intermittent reinforcement extinction phenomenon was obtained.

Basic to the thesis of Crum, Brown, and Bitterman, and similar interpretations, is the implicit notion that greater resistance to extinction is found following *variability*, due either to intermittency or variable delay of reinforcement, than following *consistency*, continuous immediate reinforcement, during acquisition. However, with respect to variability in terms of delay of reinforcement, it is possible that it is not the variability that produces the effect, but the delay per se. The only study bearing upon this question is Fehrer's (**4**). She found greater resistance to extinction following a constant 20-sec. delay of reinforcement than following immediate reinforcement. However, for purposes of her experiment, *Ss* were delayed in the goal box rather

[1] Based upon a thesis submitted to the Graduate School of the University of Wisconsin in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer is indebted to his major professor, L. B. Wyckoff, Jr., under whose direction the research was carried out. The writer was a National Science Foundation Predoctoral Fellow throughout the course of the investigation.

[2] Now at the University of Miami.

than in a separate delay chamber. It has been known since the early studies of Watson (**23**) and Warden and Haas (**22**) that delay in the goal box does not retard acquisition, and on the basis of the principle of secondary reinforcement (**8**) resistance might then be expected to be greater following delay.

These comments point to a need for further study of the effects of *constant* delay of reinforcement on resistance to extinction. Such study would aid in interpreting the effects of intermittent reinforcement on resistance to extinction. In addition, other reasons exist for the desirability of obtaining additional evidence on this relationship.

A straightforward prediction from Hullian theory (**9**), utilizing Corollary iiiB and Postulate XVI would indicate that resistance to extinction should be greater following continuous immediate reinforcement than following constant delay of reinforcement. Following a constant number of acquisition trials, reaction potential should be greater the shorter the delay of reinforcement. But, the greater the reaction potential at the onset of experimental extinction, the greater the resistance to extinction of the instrumental response associated with that reaction potential. Thus, the shorter the delay of reinforcement during acquisition, the greater should be the resistance to extinction. A prediction of greater resistance to extinction following longer delays does not follow directly from these assumptions.

Finally, evidence on the quantitative relationship between delay of reinforcement and resistance to extinction is totally lacking. Whatever the relationship, it should have relevance for any theoretical interpretation of the learning process.

The study to be reported investigates the functional relationships between delay of reinforcement on the one hand, and acquisition, response perseveration, and reversal of a position discrimination on the other. It was designed with the following objectives: (*a*) to test the assumption that resistance to extinction is greater following intermittent reinforcement than following continuous reinforcement because of the inherent variability in the former situation; (*b*) to test the implications of the Hullian formulation regarding constant delay of reinforcement and resistance to extinction; and (*c*) to obtain empirical evidence concerning functional relationships obtaining between delay of reinforcement, and acquisition, response perseveration, and discrimination reversal.

The response perseveration measure, here used in a choice situation, is analogous to the various resistance to extinction measures used in single response situations. It seems that any theoretical position which makes a statement regarding resistance to extinction would probably make a similar statement about response perseveration in the present situation. To the extent that the perseverative error measure is of the same class of measures as are resistance to extinction measures, it will have relevance to statements regarding resistance to extinction.

Method

Experimental Design

The problem was the learning of a position discrimination in a simple maze. All *Ss*—albino rats—were run a constant number of trials and then the correct position was reversed. A 4 × 4 factorial design was employed, in which each of four delays of reinforcement during acquisition was paired with four delays during

discrimination reversal. The delay intervals were 2, 4, 8, and 16 sec.

Although there were only two procedural phases of the experiment—the acquisition and reversal phases—three measures of performance were obtained, one relating to acquisition, and two relating to performance during the reversal phase. In addition to a measure of reversal learning, the measure of perseveration of the originally correct response (the analogue of resistance to extinction) was also obtained in the second phase.

Regardless of the delay of reinforcement to which they were subjected, all *S*s received a constant 80 acquisition trials before the correct position was reversed.

Analysis of the acquisition measure and the reversal measures should yield evidence regarding functional relationships, and bearing upon the Hullian view. Analysis of the perseveration scores will yield information pertaining to the intermittent reinforcement interpretation question.

Apparatus

A one-unit Y alley maze was constructed of plywood floor and pine sides, and was covered with hinged hardware cloth. The parts were a starting box, runway and choice point, two delay chambers, and two goal boxes. The entire maze was painted flat gray. Guillotine doors separated the delay chambers from the runway, and one-way doors separated the goal boxes from the delay chambers. A further description of the maze, including specifications, will be found elsewhere (**15**).

The floor of the apparatus rested on a table top 40 in. above the floor. The only source of light was a 15-w. bulb, suspended 20 in. above the maze bifurcation.

Delay intervals, the time elapsing between lowering the guillotine door and removing the rod barring entrance to the goal box, were timed manually by a Standard Electric timer.

Subjects

Subjects were 80 male albino rats of the Sprague-Dawley strain. All were between 80 and 100 days of age at the start of the experiment. As there were 16 treatment combinations, or cells, in the experiment, there were a total of five replications.

Procedure

General conditions.—Except during experimental sessions, *S*s were housed, two or three to a cage, in a room separate from the experimental room. Food was available only as reinforcement during experimental sessions, and in the home cages for 1 hr. immediately following. Water was available in the home cages at all times. The maintenance diet was Purina Chow pellets, and the reinforcement pellets, weighing approximately 160 mgm. each, were composed of 25% sugar syrup and 75% Okatie Diet food powder.

Preliminary training.—All *S*s were run in groups of eight. Thus the running of one replication required two such groups, each group requiring 1 hr. per daily experimental session. On Day 1, diurnal deprivation cycle was begun. On each of the five following days, each group was handled for about 15 min. just preceding its feeding hour.

On Day 7, each *S* received one preliminary trial in the Y maze, with the goal box doors absent, thus allowing immediate entry into the goal boxes. One pellet was available in each goal box and *S* was allowed to remain in the apparatus until one of these had been consumed. On Days 8 and 9, two trials were given each *S* under the same conditions. On the tenth and final day of preliminary training, each *S* was given five trials, with one pellet available in each goal box, and the doors over the goal box entrances present. On the basis of these last five trials, position preferences were determined.

Acquisition training.—For the acquisition phase, each *S* was given 10 trials per day for 8 days, or a total of 80 trials. During this time, a response was considered as having been made when *S*'s nose or forepaws touched the door to the goal box, at which time the sliding door between the runway and delay chamber was lowered, and simultaneously the timer was started. At the end of the appropriate delay interval, the rod blocking the goal box entrance was removed allowing *S* to push beneath the door and enter the goal box.

If *S* had made a correct response, food was found as described above. The *S* was allowed to remain in the goal box until the pellet was consumed. If *S* had made an incorrect response, he was delayed the appropriate time in the delay chamber, then allowed to enter the empty goal box, from which he was removed after 10 sec.

All rats were run to their nonpreferred side, and the noncorrection procedure was used. The intertrial interval was 4–6 min.

For the purposes of obtaining an acquisition measure, a criterion of learning was set at two successive days with at least 19 correct responses out of the possible 20 on those two days. The measure was the number of days required to attain this criterion. All *S*s achieved this

criterion well within the 80 acquisition trials administered.

Discrimination reversal training.—Beginning the next day, a discrimination reversal procedure was introduced, the formerly incorrect side of the maze now being correct. At this same time, the reversal delay schedule went into effect. In all other respects *S*s were treated as they had been during the acquisition phase.

Two measures were obtained during reversal training. The first was a measure of perseveration of the now incorrect response. This was the number of consecutive errors before the first response was made to the now correct side. This measure is independent of any effects of reinforcements for responses to the now correct side. In this sense, the measure resembles resistance to extinction measures, whereas a measure of reversal learning would not. The other measure was of reversal learning, i.e., the the number of days required to reach a criterion of two successive days with at least 19 correct responses out of the possible 20 on those two days, as measured from the first day of the reversal procedure. Training for any *S* was discontinued after completion of criterion trials.

Results

Acquisition

Days to acquisition criterion as a function of the values of delay during acquisition and during reversal are shown in Fig. 1. Although the latter variable, delay during reversal, was not yet operative, this breakdown has been included in the figure to give some indication of the comparability of the groups. It should be noted that each *S* is represented twice in this figure, once on each curve. Inspection of the figure would indicate that the effect of delay during acquisition is a negatively accelerated, increasing one, while, as expected, the effect of delay during reversal seems to be negligible.

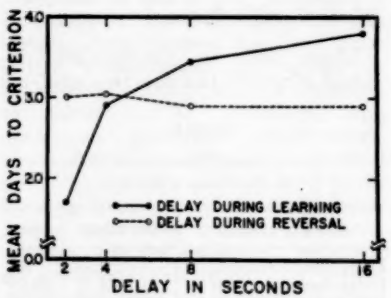

FIG. 1. Mean days to acquisition criterion as a function of delay interval during learning and reversal.

TABLE 1

ANALYSIS OF VARIANCE OF DAYS TO ACQUISITION CRITERION

Source of Variance	*df*	*MS*	*F*
Delay during learning	3	16.91	38.66*
Delay during reversal	3	.11	.26
Interaction	9	.42	.97
Within groups (error)	64	.44	

* $P = .001$.

Analysis of variance of days to acquisition criterion (Table 1) confirms the trends shown in this figure. The effect of delay of reinforcement during acquisition is highly significant, while neither of the other two possible sources of variance approaches significance. The mean days to criterion for the 2-, 4-, 8-, and 16-sec. delay during acquisition (all reversal delays combined) are 1.70, 2.90, 3.45, and 3.80, respectively.

The significant effect of delay during acquisition was further analyzed by means of the method of orthogonal polynomials applied to the relation between performance and the logarithm of the delay interval (log values are evenly spaced). This analysis revealed that both the linear and quadratic components of trend were significant (linear $F = 107.25$, $P = .001$; quadratic $F = 8.26$, $P = .01$; $df = 1$ and 64 in both cases). This significant negative acceleration in terms of log values of delay also indicates a significant negative acceleration when

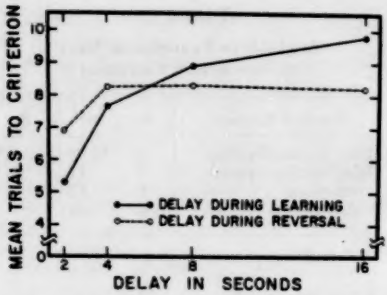

FIG. 2. Mean trials to perseveration criterion (number of perseverative errors before the first reversal response) as a function of delay interval during learning and reversal.

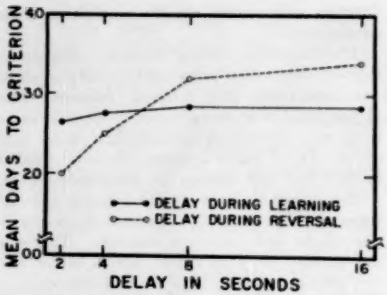

FIG. 3. Mean days to reversal criterion as a function of delay interval during learning and reversal.

the values of delay are on a linear scale.

Response Perseveration

The mean number of consecutive perseverative errors before the first reversal response as a function of delay of reinforcement during prior acquisition and delay of nonreinforcement during the perseveration trials is shown in Fig. 2. The relationship between delay during acquisition and the mean number of perseverative errors would seem to be a negatively accelerated, increasing one, such that, the greater the prior delay, the greater the perseveration. The effect of delay during reversal seems to be of the same form, but less pronounced.

TABLE 2

ANALYSIS OF VARIANCE OF RESPONSE PERSEVERATION DURING REVERSAL LEARNING

Source of Variance	*df*	*MS*	*F*
Delay during learning	3	76.21	17.88*
Delay during reversal	3	9.15	2.15
Interaction	9	3.06	.72
Within cells (error)	64	4.26	

* $P = .001$.

Analysis of variance of these data (Table 2) reveals that only the effect of delay during acquisition is significant. The mean perseverative errors following 2-, 4-, 8-, and 16-sec. delay during acquisition are 5.30, 7.65, 8.90, and 9.80, respectively.

This effect was further analyzed by the method of orthogonal polynomials. Only the linear component of trend was significant ($F = 51.05$, $P = .001$, $df = 1$ and 64). This indicates no significant deviation from a straight line for the function relating perseveration to the logarithm of the delay interval, but leaves open the question of curvature for the relation to delay itself.

Discrimination Reversal

Mean days to discrimination-reversal criterion as a function of delay

TABLE 3

ANALYSIS OF VARIANCE OF DAYS TO REVERSAL CRITERION

Source of Variance	*df*	*MS*	*F*
Delay during learning	3	.18	.46
Delay during reversal	3	8.32	20.79*
Interaction	9	.09	.24
Within cells (error)	64	.40	

* $P = .001$.

during acquisition and during reversal are shown in Fig. 3, where it will be seen that the major trend is an increase in days to criterion with increasing delays during reversal training, but not with delays during original learning.

Analysis of variance (Table 3) bears this out. Only delay during reversal proved to have a significant effect on performance. Mean days to criterion for 2-, 4-, 8-, and 16-sec. delay *during reversal* are 2.00, 2.50, 3.20, and 3.40, respectively.

Orthogonal polynomials analysis revealed that the only significant component of this trend was the linear component ($F = 60.03$, $P = .001$, $df = 1$ and 64). Again the question of curvature on a linear time scale remains unanswered.

Discussion

By the nature of the choice learning situation employed in the present experiment, it was not possible to obtain a measure of resistance to extinction of the original response. However, the response perseveration measure is considered as the analogue of resistance to extinction, both for purposes of comparison with other experiments and as bearing upon the various theoretical positions.

Results of the Crum, Brown, and Bitterman (1) and Scott and Wike (17) investigations revealed that both intermittent reinforcement and variable delay of reinforcement increase resistance to extinction, as contrasted with continuous immediate reinforcement. Crum, Brown, and Bitterman suggested implicitly that it is the variability in the learning situation, which both of the former procedures produce, that contributes to the enhanced resistance to extinction. However, the present experiment reveals that the delay alone, with no variability, produces increased resistance. There was no experimentally imposed variability in the present experiment, and yet, the greater the delay, the greater the resistance to extinction.

These findings do not negate interpretations of intermittent reinforcement extinction phenomena in terms of what happens on nonreinforced trials (1, 24). However, they do indicate that it is not necessarily the variability which produces the effect.

The acquisition results of the present study are quite in accord with the findings of previous investigations of the effects of delay of reinforcement on acquisition, and with the Hullian position, in that improvement is slower the greater the delay, and that the function is a negatively accelerated one.

However, the perseveration results are the opposite of what the Hullian formulation must predict. According to the Hullian derivation, resistance to extinction should be greater the less the prior delay. Results of the present investigation indicate just the opposite. Other interpretations within the Hullian framework are perhaps possible, such as by use of the concept of fractional antedating goal reaction; however, these possibilities will not be explored.

Another interpretation of extinction phenomena for which the results of the present investigation hold relevance is that of Skinner (20). Skinner views extinction in terms of the novelty of the extinction situation as contrasted with the acquisition situation. The more similar are acquisition and extinction conditions, the greater should be the resistance to extinction. This view has been used as an interpretation of greater resistance to extinction following intermittent reinforcement than following continuous reinforcement. In the present context, delay of reinforcement resembles extinction more than does no delay, and the greater the delay the greater the resemblance; thus resistance should be greater following longer delays during acquisition. The present results are in accord with this position.

Using this notion of similarity between acquisition and extinction conditions as a framework, it should be possible

to devise a model which will more adequately describe the results of the present study. One approach to an interpretation of both acquisition (original or reversal) and extinction results within a single framework, as a function of delay of reinforcement, is a conceptualization in which the amount of learning or extinction is *not* the same for all delay intervals. Rather, it might be conceptualized that, the shorter the delay during acquisition, the less there is to learn or extinguish.

Whereas Hull (8, 9), Guthrie (6), Estes (3), Skinner (19), and others, have considered learning to be essentially an increase in the tendency to perform some experimentally designated correct response, and extinction to be a decrease in this tendency, an alternative approach, which satisfies the above conditions, would be to view learning as a decrease in the strength of competing response tendencies (7), and extinction as an increase in strength of competing response tendencies. It is felt that an approach along these lines will better satisfy the results of the present and related experiments than would the conventional approach.[3]

It might be assumed that when there is delay of reinforcement present, competing responses occur in the time interval between the occurrence of the instrumental (choice) response and receipt of reinforcement, and that the greater the delay interval the greater the number of competing responses which will occur. Thus, the maximum number of competing responses would occur with infinite delay, which is characteristic of nonreinforcement, of which both extinction and incorrect instrumental responses during acquisition are examples.[4] If it is further assumed that some constant proportion of the competing responses adapts out if followed by reinforcement, and that a constant proportion of adapted responses de-adapt, or become actively competing if followed by nonreinforcement, the acquisition, extinction (perseveration), and reversal results follow.

The greater the delay of reinforcement during acquisition training, the greater the number of competing responses which must adapt out, and hence the greater the number of trials to a learning criterion (represented by some maximum number of competing responses remaining).

During extinction, delay is now effectively infinite for all animals, so the population of *potential* competing responses will be the same regardless of the actual time spent in the delay chamber. But the greater the prior delay, the greater the number of competing responses which will have adapted out. Hence more must de-adapt for an extinction criterion to be attained.

If a simplifying assumption is made that reversal learning does not begin until after extinction criterion has been achieved, reversal learning will follow the same course as did original learning.

Some evidence that the last assumption is justified comes from a recent study by Eisman, Asimow, and Maltzman (2), who found that drive level during original learning affected perseveration of the formerly correct response during discrimination reversal, but not trials to reversal criterion. It would seem that persistent effects of parameters of reinforcement during acquisition upon discrimination reversal are short-lived.

The results of the present investigation, as well as the model just described, are in general agreement with the studies of variable delay of reinforcement cited earlier. The only incongruity comes when the Kintsch and Wike study (10) is compared with the present one. They found that length of delay during acquisition had a significant effect upon reversal learning. However, all of their

[3] A detailed mathematical description of this model, including quantitative predictions of the obtained results, may be found in the author's Ph.D. thesis (16).

[4] It is to be noted that Spence (21, p. 153) has recently conceptualized experimental extinction in this fashion, as infinite delay of reinforcement.

Ss were reversed under a condition of constant immediate reinforcement for correct responses. Possible effects of change of delay and amount of change of delay are uncertain, and could not be evaluated in their experiment.

Summary

An investigation was undertaken to determine the functional relationships between delay of reinforcement, and acquisition, response perseveration, and discrimination reversal. Each of four delays during acquisition was paired with four delays during reversal. The delay intervals were 2, 4, 8, and 16 sec.

Eighty male albino rats were given 80 acquisition trials, 10 per day, in a one-unit Y alley maze. Following this, the formerly incorrect side was made correct. Five rats were run under each of the 16 delay conditions of the experiment.

The results may be summarized as follows:

1. The number of days to the discrimination learning criterion was found to be a negatively accelerated, increasing function of delay of reinforcement during acquisition.

2. Response perseveration—the number of perseverative errors before the first reversal response—was an increasing function of delay of reinforcement during prior acquisition. This trend is probably also negatively accelerated. A trend toward increasing perseveration with increasing delays during reversal was not significant.

3. The number of days to the discrimination reversal criterion was found to be an increasing function of delay of reinforcement during reversal training. This trend is probably also negatively accelerated.

The results were discussed in relation to interpretations of intermittent reinforcement extinction phenomena employing the notion of variability of training conditions, and the theoretical positions of Hull and Skinner. An interpretation was developed in terms of the assumptions that acquisition is brought about through the adapting out of competing responses, and extinction through de-adapting of adapted competing responses.

References

1. Crum, J., Brown, W. L., & Bitterman, M. E. The effect of partial and delayed reinforcement on resistance to extinction. *Amer. J. Psychol.*, 1951, **64,** 228–237.
2. Eisman, E., Asimow, A., & Maltzman, I. Habit strength as a function of drive in a brightness discrimination problem. *J. exp. Psychol.*, 1956, **52,** 58–64.
3. Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57,** 94–107.
4. Fehrer, E. Effects of amount of reinforcement and of pre- and postreinforcement delays on learning and extinction. *J. exp. Psychol.*, 1956, **52,** 167–176.
5. Grice, G. R. The relation of secondary reinforcement to delayed reward in visual discrimination learning. *J. exp. Psychol.*, 1948, **38,** 1–16.
6. Guthrie, E. R. *The psychology of learning.* (Rev.) New York: Harper, 1952.
7. Harlow, H. F., & Hicks, L. H. Discrimination learning theory: Uniprocess vs. duoprocess. *Psychol. Rev.*, 1957, **64,** 104–109.
8. Hull, C. L. *Principles of behavior.* New York: Appleton-Century Crofts, 1943.
9. Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
10. Kintsch, W., & Wike, E. L. Habit reversal as a function of length of partial delay of reinforcement. *Psychol. Rep.*, 1957, **3,** 11–14.
11. Logan, F. A., Beier, E. M., & Kincaid, W. D. Extinction following partial and varied reinforcement. *J. exp. Psychol.*, 1956, **52,** 65–70.
12. Perin, C. T. The effect of delayed reinforcement upon the differentiation of bar responses in white rats. *J. exp. Psychol.*, 1943, **32,** 95–109.
13. Perkins, C. C. The relation of secondary reward to gradients of reinforcement. *J. exp. Psychol.*, 1947, **37,** 377–392.
14. Peterson, L. R. Variable delayed reinforcement. *J. comp. physiol. Psychol.*, 1956, **49,** 232–234.
15. Pubols, B. H. The facilitation of visual and spatial discrimination reversal by overlearning. *J. comp. physiol. Psychol.*, 1956, **49,** 243–248.
16. Pubols, B. H. Delay of reinforcement, response perseveration, and discrimination reversal. Unpublished Ph.D. thesis, Univer. of Wisconsin, 1957.
17. Scott, E. D., & Wike E. L. The effect of partially delayed reinforcement and trial-distribution on the extinction of an instrumental response. *Amer. J. Psychol.*, 1956, **69,** 264–268.
18. Sheffield, V. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39,** 511–526.

19. SKINNER, B. F. *The behavior of organisms.* New York: Appleton-Century-Crofts, 1938.
20. SKINNER, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, **57**, 193–216.
21. SPENCE, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.
22. WARDEN, C. J., & HAAS, E. L. The effect of short intervals of delay in feeding upon speed of maze learning. *J. comp. Psychol.*, 1927, **7**, 107–116.
23. WATSON, J. B. The effect of delayed feeding upon reaction. *Psychobiology*, 1917, **1**, 51–60.
24. WEINSTOCK, S. Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *J. comp. physiol. Psychol.*, 1954, **47**, 318–322.
25. WOLFE, J. B. The effect of delayed reward upon learning in the white rat. *J. comp. Psychol.*, 1934, **17**, 1–21.

(Received August 21, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

DISCRIMINATION LEARNING AS A FUNCTION OF PRETRAINING REINFORCEMENT SCHEDULES[1]

HAROLD W. STEVENSON AND LEO A. PIROJNIKOFF

University of Texas

It has been found in numerous studies that partial reinforcement produces greater resistance to extinction than does continuous reinforcement (**2**). This result has been found primarily in studies where response strength is measured following the withdrawal of reinforcement. Two recent studies have shown that this finding may also be generalized to conditions where response strength is assessed during performance on a response reversal problem (**1, 3**). In both studies the reversal was acquired more readily after continuous than after partial reinforcement, indicating that extinction of the previously correct response occurred more slowly in the partially reinforced *S*s.

The present study attempts to determine whether the greater resistance to extinction following partial reinforcement can be obtained by a procedure in which *S*s are required to discriminate among a set of stimuli following a period in which all responses have received either continuous or partial reinforcement. This procedure differs from that of the response reversal studies. In reversal problems the response to one stimulus is reinforced, and the response to another is not. When the appropriate response has been learned, the conditions are reversed, and a response to the previously unreinforced stimulus is reinforced. In the present study the responses to all stimuli are reinforced during a pretraining period, and then the conditions are changed so that the response to only one of the stimuli is reinforced.

A group which was not reinforced during the pretraining phase was included. The inclusion of this group made it possible to contrast the effects of 0% reinforcement with continuous and 50% reinforcement.

METHOD

Subjects.—The *S*s were 48 children obtained from nursery schools in Austin, Texas.[2] They were randomly divided into four groups of 12 *S*s each. One group received 100% reinforcement during pretraining; a second received 50%, and a third received 0%; the fourth group of *S*s will be described later. The CA's of *S*s ranged from 3.5 to 4.7 yr. The groups were well matched by age, and the mean CA for each of the four groups was 4.1 yr.

Materials.—The materials consisted of (*a*) a 14 × 16-in. tray; (*b*) a screen to shield *E*; (*c*) three blocks, a circle with a 4-in. diameter, a triangle with 4.5-in. sides, and a square with 4-in. sides; (*d*) a drawing of a tree and landscape attached to the screen; and (*e*) colored stickers of flowers, birds, and animals. The board and screen were painted flat black, and the blocks were painted bright yellow.

Procedure.—The *S* was seated across from *E*. The *E* explained, "This is a game we play with stickers. See these stickers. Let's paste them on this picture." The *E* gave *S* three stickers and showed *S* how to paste them on the picture. The *S* was told that a game would be played where he could find more stickers. The screen was rotated so that the blocks on the board were exposed, and the following instructions were given: "See these blocks? [The *E* picked up the blocks one at a time.]

[1] This study was supported by a grant from the University of Texas Research Institute.

[2] The writers wish to express their appreciation to the directors of the following nursery schools in Austin, Texas, for their cooperation in making *S*s available: University of Texas, Enfield, Austin Community, and Tarrytown Methodist Nursery Schools.

I'm going to hide the sticker and see if you can find it." The *S* was told that he could keep the picture when the game was finished. Only one choice was allowed per trial. The horizontal arrangement of the blocks on the tray was random for successive trials, and the same random order was used for all *Ss*.

Three conditions were created by varying the frequency of reinforcement during the 20-trial pretraining period. In the 100% group a sticker was placed under each of the three blocks for each of the 20 trials; therefore *S* was reinforced on every trial regardless of which block he chose. In the 50% group a sticker was placed under each of the blocks in 10 of the 20 trials. The trials on which *S*'s responses were reinforced were determined at random prior to the experiment. This procedure insured that *S* would be reinforced on half of the trials, and that reinforcement would not depend upon the specific response *S* made. In the 0% group no stickers were placed under the blocks for the 20 trials; thus *Ss* received no reinforcement during pretraining.

Immediately after the 20 pretraining trials the discrimination problem was introduced without comment. The *S* was now required to discriminate the correct block on every trial. The block under which the sticker was to be found was varied at random among *Ss*, and for a single *S* the sticker was found only under this block. During these trials the position of the blocks on the tray was again varied at random.

The trials continued until four successive correct responses were made or until 40 trials had been given. It was assumed that learning had occurred if *S* made four successive correct responses. A procedure was introduced to reduce position responses. If *S* selected a block in the same position for four successive trials, the reinforced block was randomly alternated between the two nonpreferred positions until *S* changed his response. At the end of the experiment *E* gave *S* an additional number of stickers so that each *S* left with approximately the same number on his picture.

A fourth group of *Ss* (NP group) was given the discrimination problem with no pretraining trials. This group was included to determine the rate at which the discrimination could be learned with no prior experience with the stimuli.

TABLE 1

ANALYSIS OF VARIANCE

Source	*df*	*MS*	*F*
Conditions (C)	2	76.96	4.34*
Subjects (*Ss*)	33	15.42	
Trials (T)	3	37.84	19.70***
C × T	6	6.40	3.33**
Ss × T†	99	1.92	
Total	143		

* $P < .05$.
** $P < .01$.
*** $P < .001$.
† Bartlett's test, $P > .20$.

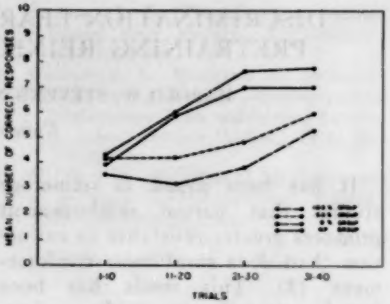

FIG. 1. Mean number of correct responses in blocks of 10 trials for each experimental group.

RESULTS

The mean numbers of correct responses made in the 100%, 50%, and 0% groups during the 40 learning trials were 26.2, 19.1, and 15.8. The mean for the group which had no pretraining was 23.8. The numbers of *Ss* learning the discrimination in the 100%, 50%, and 0% groups were 8, 5, and 3. Seven *Ss* with no pretraining learned the problem. Among the groups given pretraining the rate of learning was found to vary directly with the amount of reinforcement received during pretraining.

An analysis of variance was performed to evaluate the significance of the differences in numbers of correct responses made by *Ss* in the 100%, 50%, and 0% groups. As

seen in Table 1, the differences are significant beyond the .05 level.

The average numbers of correct responses made on successive blocks of 10 trials by each group are presented in Fig. 1. During the first 10 trials the responses were at a chance level. After this, differences among the groups became manifest and were maintained consistently throughout the remaining trials. The changes in performance during the 40 trials are highly significant (see Table 1), revealing that all of the groups showed an improvement in performance. The curves for the 100% group and the group given no pretraining are negatively accelerated, while those for the 0% and 50% groups show positive acceleration. The interaction between conditions and trials is significant beyond the .05 level. It is concluded that the learning process for the three groups differed significantly.

The relationship between the NP group and the three pretraining groups was analyzed by means of *t*. The group given no pretraining did not differ significantly in number of correct responses from the 100% group ($t = 1.36$; $P > .05$), but did differ significantly from both the 50% ($t = 2.94$; $P < .01$) and the 0% groups ($t = 3.50$; $P < .001$).

Observation of the children indicated marked differences in the behavior among the three groups during the pretraining trials. The use of no reinforcement appeared to result in great frustration. This was evidenced by such behavior as slamming the blocks and vacillating between response and withdrawal. The *Ss* in the 50% group appeared to be confused. This was seen in attempts to pick up more than one block, in restlessness, and in perturbed behavior when a sticker was not found. None of these types of behavior was characteristic of *Ss* in the 100% group, whose performance was more efficient and smooth. The different modes of response are reflected in the amount of time required to complete the 20 pretraining trials. The *Ss* in the 100%, 50%, and 0% groups required a mean of 8.7, 11.1, and 16.1 min., respectively. These scores do not differ significantly, but are in line with the differences in behavior observed.

An analysis of the responses on the pretraining trials indicated that the procedure employed was effective in producing equal response tendencies toward each of the three blocks. During the 20 trials each block was chosen approximately an equal number of times. The proportion of times that the circle, triangle, and square were chosen was .35, .34, and .31.

Discussion

The results of the study indicate striking differences in the rate at which *Ss* in the three groups given different frequencies of reinforcement during pretraining later learned the correct response in the discrimination problem.

The similarity in performance between the NP and 100% groups indicates that the pretraining received by the 100% group did not significantly retard their rate of learning. The acquisition curves for both groups show the negative acceleration commonly found in learning curves. Pretraining significantly retarded learning in the 50% and 0% groups, and the learning curves of these groups show positive rather than negative acceleration. The form of these curves may be interpreted as reflecting the persistence of inhibitory tendencies developed during pretraining. The performance of *Ss* in these groups did not begin to improve until after approximately 20 trials, and not until Trial 40 were they at the level reached by *Ss* in the 100% group in 20 trials.

Summary

Groups of *Ss* were given either 100%, 50%, or 0% reinforcement during a pretraining series of 20 trials. Following this, all *Ss* were presented with a form discrimination task involving the stimuli used in pretraining. A fourth group of *Ss* learned the discrimination with no prior experience with the stimuli. Significant differences among the first three groups in the frequency of correct response during the learning trials and a significant interaction between trials and conditions were obtained. The performance of the fourth group of *Ss* did not differ significantly from that of the 100% group.

References

1. Grosslight, J. H., Hall, J. F., & Scott, W. Reinforcement schedules in habit reversal —a confirmation. *J. exp. Psychol.*, 1954, **48**, 173–174.
2. Jenkins, W. O., & Stanley, J. C., Jr. Partial reinforcement: A review and critique. *Psychol. Bull.*, 1950, **47**, 193–234.
3. Wike, E. L. Extinction of a partially and continuously reinforced response with and without a rewarded alternative. *J. exp. Psychol.*, 1953, **46**, 255–260.

(Received August 23, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

A POSITIVE CORRELATION BETWEEN ACTION DECREMENT AND LEARNING[1]

EDWARD L. WALKER AND NOEL PARADISE

University of Michigan

In previous papers (**1, 2**) an hypothesis has been stated which asserts that the greater the temporary action decrement resulting from the performance of an act, the greater the learning that will result from that performance. Studies designed to show the effect of reward and motivation on the amount of action decrement make clear that predicted differences in the amount learned on a single trial are too small to be demonstrable without an impractical *N*.

Two studies of alternation phenomena (**3, 4**) demonstrated differences in the tendency to alternate between several conditions. One study was carried out in a horizontal maze (H) and the other in a vertical or contortion maze (V). In each maze an effort was made to analyze the contribution of the stimulus (S), the place (P), and the response (R) to the total alternation tendency (SPR).

When the two alternation studies are taken together, the capacity to produce decrement running from greatest to least was ordered SPR-V > SPR-H > S-H > S-V = R-V > P-H > R-H > P-V. The hypothesis tested in this study is that the rate of learning in eight similar conditions will be ordered in the same way.

METHOD

Subjects.—The *S*s were 64 albino rats. They were experimentally naive and were approximately 90 days of age. They were purchased from a local laboratory supplier.

Apparatus.—The original alternation study (**3**) had involved two T mazes with common arms but different stems. All parts of the floor were at the same level. The second study (**4**) employed a maze as nearly identical as possible except that the choice point twisted in three dimensions, and the starting stem was thus lower than the remainder of the maze. These two studies were run in different but very similar rooms. For the present study a new horizontal choice point was built that could be used in the maze with the twisted choice point. All conditions of this study were run in the same room and the same maze with alternative choice points—horizontal and vertical.

Each version of this maze was always used as a T maze by manipulating appropriate doors. The enclosed alleys were 4 × 4 in. and starting stem and goal stems were 18 in. Floors and roof were hardware cloth. One goal arm and box was painted flat black and the other white. All else was gray. The room was small and had walls which differed in texture and color; the single fluorescent light was not directly over the maze.

Procedure.—The *S*s were placed on a 22-hr. water deprivation schedule throughout the period of pretraining and training. Pretraining consisted of four days of handling and running down a straight runway to water. The training procedure consisted of a repetition of each of the eight conditions of the alternation studies except that each was now converted into a learning problem. Within each of the two maze conditions, horizontal and vertical, there were four different learning conditions: (*a*) Groups S-H and S-V were required to select a stimulus (black or white with place and response essentially randomized; (*b*) Groups P-H and P-V learned to go to a place with stimulus and response randomized; (*c*) Groups R-H and R-V learned to turn consistently in one direction with place and stimulus randomized; and (*d*) Group SPR-H and SPR-V found the reward always associated with the same stimulus, place, and response.

In each condition there were two alternatives which might be correct, two starting stems which could be used, and two positions within the room from which the start (the maze was on

[1] This study was carried out on a grant to the senior author from the National Science Foundation.

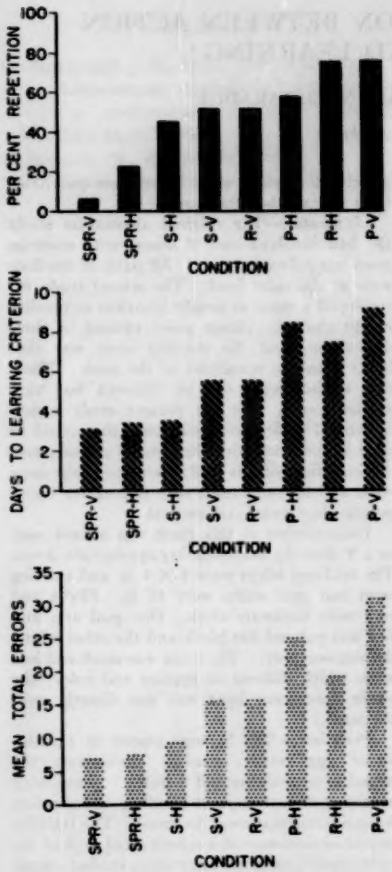

FIG. 1. A comparison of alternation scores with learning scores under similar conditions. The top figure shows the relative tendencies to alternate (expressed here as a tendency to repeat) in eight different conditions (see text for explanation of conditions). The middle figure shows the number of days required to reach a learning criterion in the same eight conditions now made a basis for learning. The bottom figure shows the mean total errors in the same eight learning conditions.

rollers and could be rotated through 180°). Therefore, eight *S*s were required in each learning group in order to control all such related preferences.

Eight trials were run each day for each *S* until a criterion of 15 correct responses in two consecutive days was reached. Massed training was used. On each trial *S* was retained in the goal box for 30 sec. and was held in the hand while the maze gates were reopened. The maze rotated if required. The time between trials was thus approximately 15 sec. under all conditions.

RESULTS AND DISCUSSION

The results may be seen in Fig. 1. The histograms at the top represent the results of the alternation studies (**3, 4**) expressed here as the relative tendency to repeat (the reverse of alternation). The other two histograms show results for the eight learning groups. The degree of agreement is immediately apparent. Kendall's Tau was run between the order in the alternation studies and the days to criterion performance in the learning studies. Tau is .89 with a probability of chance occurrence of less than .0009.

To the extent to which they are relevant, these results are taken as support for the basic hypothesis under consideration. A large tendency to alternate and thus a small tendency to repeat in an alternation situation is used as a basis for the inference that there is a large decrement in the relevant action system. Therefore, the eight conditions in the top histogram of Fig. 1 are taken to indicate different degrees of action decrement. If the amount of action decrement is positively correlated with the amount of habit strength, then the eight conditions should differ with respect to the number of trials required to reach a criterion learning performance. That is, the greater the tendency to alternate, the more the action decrement, the greater the habit strength acquired on a single trial, and thus the fewer the number of trials required to reach a criterion.

SUMMARY

An empirical test was made of the hypothesis that the greater the action decrement demonstrable in a given situation, the greater the

potential habit strength and thus the faster the learning. Groups of eight rats each were run in eight different learning problems in which the relevant cues had previously been demonstrated to have different capacities to produce action decrement. The hypothesis was confirmed to the extent that there was a highly significant relation between the amount of decrement previously demonstrated and the rate of learning.

References

1. Walker, E. L. The course and duration of the reaction decrement and the influence of reward. *J. comp. physiol. Psychol.*, 1956, **49**, 167–176.
2. Walker, E. L. Action decrement and its relation to learning. *Psychol. Rev.*, 1958, **65**, 129–142.
3. Walker, E. L., Dember, W. N., Earl, R. W., & Karoly, A. J. Choice alternation: I. Stimulus vs. place vs. response. *J. comp. physiol. Psychol.*, 1955, **48**, 19–23.
4. Walker, E. L., Dember, W. N., Earl, R. W., Fawl, C. L., & Karoly, A. J. Choice alternation: III. Response intensity vs. response discriminability. *J. comp. physiol. Psychol.*, 1955, **48**, 80–85.

(Received September 3, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

AMOUNT AND PERCENTAGE OF REINFORCEMENT AND DURATION OF GOAL CONFINEMENT IN CONDITIONING AND EXTINCTION[1]

STEWART H. HULSE, JR[2]

Brown University

This research was designed to determine the effect of amount of reinforcement and duration of goal-box confinement on the greater resistance to extinction following partial as compared with continuous reinforcement.

While several studies have been concerned with the role of amount of reinforcement in training performance when partial reinforcement has also been used (e.g., **3, 5, 6**), no study, apparently, has investigated extinction performance under such conditions. Some interesting implications can be drawn from the literature with regard to extinction performance, however. Thus, with spaced training trials and small food rewards (ca. .08 gm.) several alley studies (e.g., **9, 10, 15**) have found that the greater resistance to extinction typically produced by partial as compared with continuous reinforcement is either not obtained or can only be demonstrated by using sensitive covariance analyses. One study (**14**), however, used spaced training and a larger reward (.4 to .5 gm.) and clearly demonstrated the typical partial-reinforcement effect.

The above experiments were interpreted in terms of the effects of spaced training which, according to Sheffield's (**10**) deduction in terms of generalization decrement, should reduce the magnitude of the partial-reinforcement effect. Taken as a whole, however, they suggest that amount of reinforcement may be an important variable in obtaining the partial-reinforcement effect totally aside from spaced training per se.

Duration of goal-box confinement as a variable involved in partial-reinforcement situations has been investigated by Stanley and Clayton (**11**) in their test of Weinstock's (**14**) habituation of competing response theory concerning the partial-reinforcement effect. Stanley and Clayton inferred from the theory that long as compared with short goal-box confinements on nonfood-reinforced training trials for partially reinforced *Ss* should increase resistance to extinction. Their results do not support this deduction. Long vs. short confinements did not produce significant differences in extinction performance. The only reliable factor was whether or not goal-box confinement was *changed* from training to extinction. If a change occurred (e.g., if *Ss* were trained with a short confinement on nonreinforced trials and then extinguished with a long confinement), resistance to extinction was less than if no such change occurred.

As Stanley and Clayton note, however, the interpretation of their data as evidence against the Weinstock theory is complicated by the fact that massed trials (15-sec. intertrial interval) were used. This procedure may have introduced confusing fac-

[1] This research represents a portion of a dissertation submitted to the Graduate School of Brown University in partial fulfillment of the requirements for the Ph.D. degree. The author is greatly indebted to Walter C. Stanley who directed the dissertation.

[2] Now at Johns Hopkins University.

tors such as generalized "excitement" carrying over from trial to trial. It may be, as Weinstock implies, that his theory applies only to situations involving widely spaced trials.

The present experiment provided information concerning the above by using two amounts of reinforcement, two percentages of reinforcement, two durations of goal-box confinement on nonreinforced training trials, and two durations of goal-box confinement on extinction trials. A one-trial-per-day procedure was used throughout training and extinction.

Method

Subjects

The *S*s were 108 experimentally naive male albino rats of the Sprague-Dawley strain obtained from the Charles River Breeding Laboratories, Brookline, Massachusetts. The *S*s ranged in age from 60 to 85 days at the beginning of taming. Nine *S*s were randomly assigned to each of the 12 experimental groups.

Apparatus

The apparatus consisted of a straight enclosed alley 48 in. long and 4 in. wide. The start box was 6 in. wide and 7 in. long; the goal box was 8 in. wide and 12 in. long. Inside height throughout was 9 in. The start box and goal box had hinged glass lids, and the alley was covered by a hinged copper-screened frame. A metal food cup was placed in a far corner of the goal box. A black curtain hung diagonally across the goal box to prevent *S* from seeing the food cup until he had actually run into the goal box.

One guillotine door was located at the entrance to the alley and one at the end of the alley, 4 in. from the goal box. These doors were controlled by cords operated by *E*. The entire apparatus was painted flat black with the exception of the start box and the food cup which were painted white.

The apparatus was lighted by four 15-w. bulbs suspended 2 ft. above the alley. One of these bulbs was directly over the start box, another over the goal box, and the remaining two over the alley.

To provide time measures of *S*s' performance, as described below, microswitches were attached to the apparatus so they could be operated by movements of the start-box door and a floor section at the entrance to the goal box; a polyethylene leaf switch was also placed under a moveable floor section in the start box. The switches were connected, via appropriate relay circuits, to two Standard Electric timers.

An interval timer, consisting of a 10-sec. siliconetic relay circuit which activated a neon pilot light, was used to signal to *E* the duration of a 10-sec. goal-box confinement period. A 60-sec. siliconetic relay circuit signaled the duration of a 60-sec. goal-box confinement period. Timing of both these intervals began automatically when *S* depressed the goal-box floor.

Experimental Design

The experimental design included two levels of four variables: percentage of reinforcement (46% or 100%), amount of reinforcement (.08 gm. or 1.0 gm.), duration of goal-box confinement on nonfood-reinforced training trials (10 sec. or 60 sec.), and duration of goal-box confinement on extinction trials (10 sec. or 60 sec.). Table 1 summarizes the design, which required 12 groups of *S*s.

It will be noted in Table 1 that the design is subdivided into two units of eight groups each (Units A and B). Unit A includes two 100% groups and two 46% groups which were confined in the goal box for approximately 60 sec. on all training trials. Similarly, two 100% groups and two 46% groups were confined in the goal box for approximately 10 sec. on all training trials. This equating of goal-box confinement times was done by carrying out a pilot study to determine the amount of food that *S*s could eat in 60 sec. and 10 sec., the times chosen for goal-box confinement on nonreinforced training trials for the partially reinforced groups. Thus, the two 100% groups given a large amount of reward required approximately 60 sec. to eat. Two corresponding 46% groups given the same large reward on reinforced trials were confined for 60 sec. in the empty goal box on nonreinforced trials. The same equating of time spent in the goal box held for the two 100% and the two 46% groups given a small reward. Here, eating times were just under 10 sec. and confinement time on nonreinforced trials for the 46% groups was 10 sec. By thus equating time spent in the goal box for these eight groups, the design permits determining the effects on performance of (*a*) percentage of reinforcement, (*b*) amount of reinforcement, and (*c*) extinction confinement time.

TABLE 1

EXPERIMENTAL DESIGN

($N = 9$ in each group)

Analytic Unit	Group Designation*	Percentage of Reinforcement	Amount of Reinforcement	Duration of Goal-Box Confinement	
				Training	Extinction
A	100Lg-L	100%	1.0 gm.	—	60 sec.
A	100Lg-S	100%	1.0 gm.	—	10 sec.
A	100Sm-L	100%	.08 gm.	—	60 sec.
A	100Sm-S	100%	.08 gm.	—	10 sec.
B	46LgLL	46%	1.0 gm.	60 sec.	60 sec.
B	46LgLS	46%	1.0 gm.	60 sec.	10 sec.
B	46SmSL	46%	.08 gm.	10 sec.	60 sec.
B	46SmSS	46%	.08 gm.	10 sec.	10 sec.
B	46LgSL	46%	1.0 gm.	10 sec.	60 sec.
B	46LgSS	46%	1.0 gm.	10 sec.	10 sec.
B	46SmLL	46%	.08 gm.	60 sec.	60 sec.
B	46SmLS	46%	.08 gm.	60 sec.	10 sec.

* The group designations summarize the experimental treatment for each group, where 100 and 46 show percentage of reinforcement, Lg and Sm show 1.0 and .08 gm. rewards, and L and S show 60-sec. and 10-sec. confinement times. The first L or S refers to confinement on nonreinforced training trials; the second refers to confinement during extinction. Only the second L or S appears for the 100 groups since they received no nonreinforced training trials.

Unit B includes the eight 46% groups. By thus holding percentage of reinforcement constant at 46%, the design permits determining the effects on performance of (*a*) goal-box confinement time on nonreinforced training trials, (*b*) extinction confinement time, and (*c*) amount of reinforcement.

Procedure

Taming.—Each *S* was tamed for two weeks, receiving as its daily ration of wet mash 7.5 gm. (dry weight) of Purina Lab Chow. During the first week, each *S* was fed ½ of its daily ration in each of three individual cages. The *S* was freely handled as it was transferred from cage to cage. During the second week, *S*s were also allowed to explore in batches of four in a larger wooden enclosure. Throughout taming and the remainder of the experiment, *S*s were housed in colony cages, four to a cage.

Pretraining.—All *S*s were given six pretraining trials, one trial per day, with 100% reinforcement. The amount variable (Table 1) was introduced immediately.

Each day, *S*s were carried four at a time in their living cages to the experimental room. They were then transferred to individual retaining cages, each *S* being assigned the same cage each day.

After 30 sec. the first *S* was placed in the start box and 2–3 sec. later the start-box door was raised. The goal-box door was lowered as soon as *S* depressed the floor section in the goal box which recorded total running time. The *S* was confined in the goal box until it had finished the food reward and was then immediately removed. Following running, each *S* was returned to its retaining cage, and after all four *S*s in a batch had been run, they were returned to their living cage and removed to the colony room.

On the first four trials the goal-box curtain was hung entirely to one side of the goal box. On Trial 5 and all training and extinction trials thereafter, the curtain was hung so that it shielded all but 2 in. of the goal box.

Food reinforcements consisted of wet mash pellets weighing either 1.0 gm. or .08 gm., dry weight. The .08-gm. pellet usually required just under 10 sec. to eat. The 1.0-gm. pellet required a median eating time for all *S*s of 68 sec., just over the time predicted from the pilot study.

The *S*s were given a supplementary feeding of wet mash in individual wooden cages, 45 to 90 min. after they had been run. This was sufficient to bring the daily ration to 7.5 gm., dry weight. The *S*s were 22–23 hr. hungry at the time trials were started each day.

None of the *S*s met a discard criterion, set in advance, of failing to begin eating the food pellet within 5 min. of the start of any one trial. Six *S*s who died during the course of the experiment were replaced by random assignment of *S*s tamed for such replacement purposes.

Training.—Following the six days of pretraining, 24 training trials were given, one trial per day, during which the variables of percentage of reinforcement and goal-box confinement time (Table 1) were introduced. On nonreinforced trials for the 46% groups, the food cup was removed from the goal box.

The 46% groups were reinforced with food on 11 of the 24 training trials. The assignment of reinforced trials was random with the following restrictions: (*a*) Trial 1 was always reinforced; (*b*) Trial 2 was never reinforced; (*c*) Trial 24 was always reinforced; and (*d*) no more than four nonreinforced trials occurred in sequence. Separate schedules of reinforcement were constructed for each of the nine replications used in the experiment.

Except for the introduction of these changes, the procedure during training was the same as that followed on the last two days of pretraining.

Extinction.—Nineteen extinction trials were given, one trial per day, following the last training trial. Confinement times were either 10 sec. or 60 sec. for any group. Food was omitted on all trials, and the food cup was removed from the goal box.

A 2-min. nonresponse criterion was used. If *S* failed to reach the goal box within 2 min., he was removed from the apparatus and returned to his retaining cage. The *S* was run again the next day. Less than 10 nonresponses were recorded for all *S*s on all 19 extinction trials. No more than 2 nonresponses were made by any one *S*.

Measures of Performance

Three measures of performance were obtained: start time, alley time, and total running time. The statistical analyses showed these three measures to be very highly correlated in measuring the effects of the experimental variables on performance. Accordingly, only the analysis of the total running time data will be reported here. Total running time included start time and alley time; it began when the start-box door went up and ended when *S* depressed the floor section in the goal box.

The response times were recorded to the nearest .1 sec., transformed to reciprocals, and multiplied by 100. Extinction times which met the 2-min. nonresponse criterion were assigned a reciprocal score of .8. Means of the reciprocal scores were determined by *S* by blocks of trials for the statistical analyses.

Results

Pretraining.—During the six pretraining trials and the first training trial, the groups were treated differently only through the use of large or small rewards. Large rewards produced significantly greater mean total running time speeds than small rewards over these seven trials ($t = 3.95$, $P < .01$ for 106 df).

Training and extinction.—During training there were six different experimental treatments; the extinction confinement variable was not yet introduced. Hence, the analysis of the training data was based on six groups with $N = 18$ in each group. The last eight training trials and the first extinction trial were used as an index of training performance. The statistical treatment of the extinction data was similar to that used for training, except that the additional variable of extinction confinement was involved. The first extinction trial was not included in the analysis.

A first type of analysis (cf. Table 1, Unit A) considered the 100% groups and the 46% groups which were equated in terms of goal-box confinement during training. This analysis permitted comparison of differences in training performance due to percentage of reinforcement and amount of reinforcement. It permitted comparison of differences in extinction performance due to percentage of reinforcement, amount of reinforcement, and extinction confinement. The results of the analyses of variance used on these groups are presented in Table 2.

Table 2 shows that, during training, both Amount of reinforcement and Percentage of reinforcement are significant variables. Large rewards produced faster mean response speeds than small rewards ($P < .01$), and continuous reinforcement produced faster mean response speeds than partial reinforcement ($P < .05$).

TABLE 2

ANALYSES OF VARIANCE FOR TRAINING TRIALS 23–31 AND FOR EXTINCTION TRIALS 2–19
(The 100% groups and the 46% groups equated for training confinement are taken as an analytic unit)

Source	Training			Extinction		
	df	*MS*	*F*	*df*	*MS*	*F*
A. Percentage of reinforcement	1	1902.42	6.48*	1	5619.77	44.04**
B. Amount of reinforcement	1	5271.93	17.96**	1	466.65	3.66
C. Extinction confinement	—	—	—	1	37.41	0.29
A × B	1	22.34	.08	1	3481.95	27.29**
A × C	—	—	—	1	250.51	1.96
B × C	—	—	—	1	696.27	5.46*
A × B × C	—	—	—	1	304.63	2.39
Within	68	293.53		64	127.61	

* $.05 > P > .01$.
** $P < .01$.

The interaction between these two variables is not significant ($P > .05$).

Table 2 also shows that, during extinction, the following are significant sources of variance: (*a*) Percentage of reinforcement ($P < .01$), (*b*) Percentage of reinforcement × Amount of reinforcement ($P < .01$), and (*c*) Amount of reinforcement × Extinction confinement ($P < .05$). Extinction confinement per se is not significant, nor are any of the other interactions. The significant interaction between Amount of reinforcement and Extinction confinement probably represents an interaction between Training confinement and Extinction confinement to be described below for the eight 46% groups. This follows since amount of reinforcement for the present eight groups was perfectly correlated with time in the goal box during training.

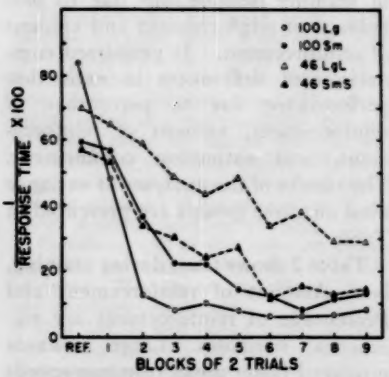

FIG. 1. Resistance to extinction of the 100% groups and the 46% groups equated for training confinement time. The reference point includes the last training trial and the first extinction trial.

Figure 1 shows the extinction performance of the four 100% and the four 46% groups in Unit A pooled for graphical presentation in terms of large and small amounts of reinforcement (empty and filled points), and 46% and 100% reinforcement (triangles and circles). The pooling is such that the nonsignificant extinction confinement variable is not represented in the curves. Table 3 shows the results of *t* tests on these pooled groups for

the variables shown to be significant by the analysis of variance.

Figure 1 and Table 3 show: (*a*) with large rewards, partial as compared with continuous reinforcement produced significantly greater resistance to extinction ($P < .01$); (*b*) with small rewards, partial as compared with continuous reinforcement produced a slight but statistically nonsignificant superiority in resistance to extinction; (*c*) with partial reinforcement, large as compared with small rewards produced significantly greater resistance to extinction ($P < .01$); and (*d*) with continuous reinforcement, large as compared with small rewards produced significantly *less* resistance to extinction ($P = .03$).

An analysis of covariance was also run on the eight groups in Unit A. This was done to examine extinction performance when differences in performance at the end of training, as measured by the last eight training trials and the first extinction trial, were statistically removed. The results of the analysis corroborate the analysis of variance in showing Percentage of reinforcement and the interaction between Percentage of reinforcement and Amount of reinforcement to be significant effects ($P < .01$). However, the interaction between Amount of reinforcement and Extinction confinement is no longer significant. The analysis of covariance also shows the triple interaction to be significant ($P < .05$); the exact meaning to be given this is not clear with the present design.

The *t* tests of the adjusted means of the groups show the same effects as does Table 3 for the unadjusted means, except that a significant partial reinforcement effect is obtained with *both* large and small rewards ($P < .01$).

TABLE 3

RESULTS OF *t* TESTS ON EXTINCTION MEAN RESPONSE SPEEDS FOR THE 100LG, 100SM, 46LGL, AND 46SMS GROUPS

Groups		Mean$_1$	Mean$_2$	*t*
1	2			
46LgL	100Lg	44.0	12.5	7.84**
46SmS	100Sm	25.0	21.3	0.92
100Sm	100Lg	21.3	12.5	2.19*
46LgL	46SmS	44.0	25.0	4.73**

Note.—The *within* error term from the analysis of variance (127.61) was used to compute *t*.
* $P = .03$, both tails and 64 *df*.
** $P < .01$, both tails and 64 *df*.

A second type of analysis (cf. Table 1, Unit B) considered the eight 46% groups. This analysis permitted comparison of differences in training performance due to amount of reinforcement and to duration of goal-box confinement on nonreinforced training trials. It permitted comparison of differences in extinction performance due to amount of reinforcement, training confinement, and extinction confinement. The results of the analyses of variance used on these groups are presented in Table 4.

Table 4 shows that, during training, only Amount of reinforcement is significant. Large rewards produced faster mean response speeds than small rewards ($P < .01$). Neither Training confinement, nor the interaction between Training confinement and Amount of reinforcement is significant.

Table 4 also shows that, during extinction, the following are significant sources of variance: (*a*) Amount of reinforcement ($P < .01$), and (*b*) Training confinement × Extinction confinement ($P < .05$). Neither Training confinement nor Extinction confinement per se are

TABLE 4

ANALYSES OF VARIANCE FOR TRAINING TRIALS 23–31 AND EXTINCTION TRIALS 2–19 TAKING THE 46% GROUPS AS AN ANALYTIC UNIT

Source	Training			Extinction		
	df	*MS*	*F*	*df*	*MS*	*F*
A. Amount of reinforcement	1	6599.09	20.40**	1	4815.07	26.89**
B. Training confinement	1	178.29	.55	1	125.93	.70
C. Extinction confinement	—	—	—	1	11.36	.06
A × B	1	52.53	.16	1	136.13	.76
A × C	—	—	—	1	156.05	.87
B × C	—	—	—	1	982.72	5.49*
A × B × C	—	—	—	1	40.15	.22
Within	68	323.50		64	179.08	

* $.05 > P > .01$.
** $P < .01$.

significant sources of variance, nor are any other interactions.

Figure 2 shows the performance of the eight 46% groups combined to demonstrate the significant interaction between training confinement and extinction confinement. If a *change* in confinement time was introduced from training to extinction (triangles), resistance to extinction was less than if no such change was introduced (circles). The group differences are very small in absolute magnitude, however, and *t* tests show none of the group comparisons to be significant.

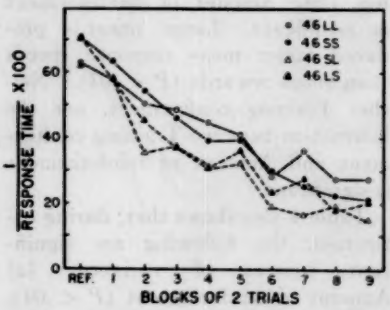

FIG. 2. Resistance to extinction of the 46% groups combined to show the effects of a change in confinement time from training to extinction. The reference point includes the last training trial and the first extinction trial.

DISCUSSION

Training.—The present data are in line with other alley studies which have shown that large rewards produce faster running during training than small rewards (e.g., **1, 8, 12, 16**), and that continuous reinforcement produces faster running during training than partial reinforcement (**7**). The fact that goal-box confinement on nonreinforced training trials was not a significant factor in affecting training performance supports, in a one-trial-per-day situation, the findings of Stanley and Clayton (**11**) with massed trials.

Extinction.—The effects of goal-box confinement on extinction performance were small in absolute magnitude compared with the effects of some of the other variables (cf. Figs. 1 and 2). The one significant effect showed that resistance to extinction was less if a *change* in confinement time occurred

from training to extinction. This duplicates, with spaced trials, the findings of Stanley and Clayton (11) with massed trials.

Different durations of training confinement did not significantly affect extinction performance. Hence, neither the present experiment (with spaced trials), nor the Stanley and Clayton experiment (with massed trials) support the deduction from the Weinstock (14) theory which predicts that long vs. short confinements on nonreinforced training trials should increase the opportunity for the habituation of competing responses, and thus increase the magnitude of the partial-reinforcement effect.

Finally, different durations of extinction confinement did not significantly affect extinction performance after either partially or continuously reinforced training. This is not in line with previous findings with continuous reinforcement (12, 13), which indicate that long extinction confinements reduce resistance to extinction. Previous studies differ from the present experiment, however, in that either more extreme confinement intervals were used (12), or massed trials were used (13). Either or both of these factors may account for the difference in results. With regard to partial reinforcement situations, the present data regarding extinction confinement agree with those of Stanley and Clayton (11).

Percentage of reinforcement and amount of reinforcement are interrelated in determining the partial-reinforcement effect. There was a significant partial-reinforcement effect of substantial proportions only when *large* rewards were used during training. Although the analysis of covariance showed a significant partial-reinforcement effect regardless of the amount of reinforcement used during training, it also showed, as did the analysis of variance, that the interaction between amount of reinforcement and percentage of reinforcement was significant. Hence, the *magnitude* of the partial reinforcement effect is clearly much greater if large rather than small rewards are used.

Besides its purely empirical interest, this finding is highly important in that it may account for discrepancies arising from tests of Sheffield's (10) deduction from the generalization decrement explanation of the partial reinforcement effect. This deduction predicts that with massed training the partial-reinforcement effect should be clearly obtained, but with spaced training it should be reduced in magnitude. With spaced training and *small* rewards, several studies (9, 10, 15) support Sheffield's deduction in that no partial reinforcement effect was obtained or else a sensitive analysis of covariance was required to demonstrate its significance. With spaced training trials and *large* rewards, however, Weinstock (14) clearly demonstrated a partial-reinforcement effect. This finding, apparently, does not support Sheffield's deduction. The present data, however, clearly suggest that the crucial variable is not the spacing of training trials per se. Rather, it is the spacing of training trials *in conjunction* with an amount of reinforcement variable which determines the magnitude of the partial reinforcement effect.

This is not the whole story. With *massed* training trials (15-sec. intertrial interval), two alley experiments (4, 10) clearly showed that a significant partial reinforcement effect could be obtained with *small* rewards (.08 gm.). No alley experiment has investigated the partial reinforcement effect with massed trials and large rewards. This is probably true because the use of large food rewards requires the spacing of trials if drive is to be held constant.[3]

The implication of the foregoing is clear: intertrial interval and amount of reinforcement *interact* in producing the partial-reinforcement effect. It will be

[3] This problem might be avoided in future research by using different concentrations of saccharine solutions as opposed to different weights of food as reinforcing agents.

clearly obtained if (*a*) large rewards and spaced trials are used, or if (*b*) small rewards and massed trials are used. As noted, it remains to be determined if the partial reinforcement effect will also be obtained with large rewards and massed trials.

A final important aspect of the data is that with continuous reinforcement, large rewards produced *less* resistance to extinction than small rewards. This is not in line with Zeaman's (**16**) findings, but is in line with the findings of others (**2, 8, 12**). The fact that Zeaman used massed extinction trials following spaced training trials may account for this difference.

Summary

This experiment investigated the role of percentage of reinforcement (100% or 46%), amount of reinforcement (1.0 gm. or .08 gm.), duration of goal-box confinement on nonreinforced training trials (10 sec. or 60 sec.), and duration of goal-box confinement on extinction trials (10 sec. or 60 sec.) as parameters of a running response. The *Ss* were 108 albino rats who were given one trial per day in a 4-ft. straight alley. Six pretraining, 24 training, and 19 extinction trials were given.

The training results show: (*a*) continuous reinforcement produced faster running than partial reinforcement ($P < .05$); (*b*) large rewards produced faster running than small rewards ($P < .01$); and (*c*) duration of training confinement did not significantly affect training performance ($P > .05$).

The extinction results show: (*a*) partial reinforcement produced greater resistance to extinction than continuous reinforcement, but the differences were much greater if large as compared with small rewards were used during training; (*b*) large as compared with small rewards produced greater resistance to extinction if partial reinforcement were used, but *less* resistance to extinction if continuous reinforcement were used; (*c*) if goal-box confinement times were *changed* from training to extinction for the partial groups, resistance to extinction was less than if no such change occurred.

The results were discussed in terms of their implications for theories which account for the fact that partial reinforcement produces greater resistance to extinction than continuous reinforcement.

References

1. Crespi, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, **55**, 467–517.
2. Fehrer, E. Effects of amount of reinforcement and of pre- and postreinforcement delays on learning and extinction. *J. exp. Psychol.*, 1956, **52**, 167–176.
3. Guttman, N. Operant conditioning, extinction and periodic reinforcement in relation to concentration of sucrose used as reinforcing agent. *J. exp. Psychol.*, 1953, **46**, 213–224.
4. Hulse, S. H., Jr., & Stanley, W. C. Extinction by omission of food as related to partial and secondary reinforcement. *J. exp. Psychol.*, 1956, **52**, 221–227.
5. Hutt, P. J. Rate of bar pressing as a function of quality and quantity of food reward. *J. comp. physiol. Psychol.*, 1954, **47**, 235–239.
6. Jenkins, W. O., & Clayton, F. L. Rate of responding and amount of reinforcement. *J. comp. physiol. Psychol.*, 1949, **42**, 174–181.
7. Jenkins, W. O., & Stanley, J. C., Jr. Partial reinforcement: A review and critique. *Psychol. Bull.*, 1950, **47**, 193–234.
8. Kling, J. W. Speed of running as a function of goal-box behavior. *J. comp. physiol. Psychol.*, 1956, **49**, 474–476.
9. Lewis, D. J. Acquisition, extinction and spontaneous recovery as a function of percentage of reinforcement and intertrial intervals. *J. exp. Psychol.*, 1956, **51**, 45–53.
10. Sheffield, V. F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39**, 511–526.
11. Stanley, W. C., & Clayton, F. L. An experimental test of Weinstock's hypothesis concerning the effect of partial reinforcement on extinction. *Psychol. Rep.*, 1955, **1**, 421–424.
12. Stanley, W. C., & Garziano, J. F. Effect of amount of reward and duration of goal confinement on conditioning and extinction. Paper read at Eastern Psychol. Ass., Philadelphia, April, 1955.
13. Stanley, W. C., & Rowe, M. I. Extinction by omission of food as a function

of goal-box confinement. *J. exp. Psychol.*, 1954, **48**, 271–274.

14. Weinstock, S. Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *J. comp. physiol. Psychol.*, 1954, **47**, 318–322.
15. Wilson, W., Weiss, E. J., & Amsel, A. Two tests of the Sheffield hypothesis concerning resistance to extinction, partial reinforcement, and distribution of practice. *J. exp. Psychol.*, 1955, **50**, 51–60.
16. Zeaman, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, **39**, 466–483.

(Received September 3, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

COMPLEX INCIDENTAL LEARNING AS A FUNCTION OF ANXIETY AND TASK DIFFICULTY[1]

CHARLES D. SPIELBERGER
Duke University

LEONARD D. GOODSTEIN
State University of Iowa

AND W. GRANT DAHLSTROM
University of North Carolina

A number of recent studies have demonstrated that the effects of motivation on performance vary with the complexity of the task involved (**3, 9, 16**). In these investigations, the Taylor Manifest Anxiety Scale (**14**), or MAS, has been utilized for selecting *Ss* who were assumed to differ systematically in drive level (**1, 2, 11**). Subjects who scored high on the MAS performed better on simple tasks and relatively more poorly on complex tasks than *Ss* who scored low. On those simple tasks (e.g., classical eyelid conditioning, paired-associate learning) in which there was but a single or highly dominant response tendency, increased drive resulted in the enhancement or facilitation of performance (e.g., **12, 15**). In more complex or difficult tasks (e.g., serial verbal learning, stylus maze learning), in which competing incorrect responses were stronger or more frequently elicited initially, or in which several competing response tendencies were aroused, increased drive has been reported to impede performance (**3, 9**). All of these findings are consistent with Hullian theory (**6**) which postulates a multiplicative relation between all habit tendencies activated in a given situation and the total effective drive then operating.

The present study is an empirical test of the relationship between drive level, operationally defined by MAS scores, and performance on the recall of a graded series of nine visual-motor tasks when no instructions to learn were given. In an earlier study (**4**) with a similar population of *Ss* it was shown that these tasks varied significantly in difficulty. Over 80% of the *Ss* recalled the easiest tasks and fewer than 40% of the *Ss* recalled the most difficult. While an interactive relationship between task difficulty and anxiety level is hypothesized, more explicit prediction would require precise knowledge of the strengths of the correct and competing responses operative for each task. Spence has recently pointed out (**12**) that with complex tasks such as these, the required knowledge of the habit and inhibitory factors operating is generally not available.

METHOD

Subjects.—The *Ss* were 88 student volunteers from introductory psychology courses at Duke University ($N = 30$), the State University of Iowa ($N = 40$), and the University of North Carolina ($N = 18$) who obtained extreme scores on the MAS. Forty-four *Ss* scored 19 or higher (HiA) and 44 *Ss* scored seven or lower (LoA) on the MAS; the High and Low groups each consisted of equal numbers of males and females.

Procedure.—Each of the designs of the Bender-Gestalt (B-G) Test was presented

[1] This research was supported in part by a grant from the Smith Fund of the University of North Carolina. The authors wish to acknowledge their indebtedness to the several research assistants at their respective institutions who aided in the collection and processing of the data.

singly to each *S* in an individual testing session. The order of presentation, however, deviated from the conventional administration of the test in such a manner as to maximize the range of task difficulty. In the present study task difficulty was defined in terms of the combined effects of both task difficulty and serial position. Task difficulty denotes the difficulty of recall of a particular design in a particular serial position as empirically determined in the previous study (4) in which the recall of these designs was found to be an interactive function of *both* serial position and design difficulty.

Each *S* was given a pencil and a pad of 4 × 6-in. white paper. The designs were presented in the order (3, 7, 4, 6, 8, 5, 2, A, 1) which placed the easiest designs in the easiest serial positions and the hardest designs in the most difficult serial positions. All *S*s were given the designs in this same order. The *S*s were not instructed to learn, but were told only that the designs would be presented individually and that each design should be copied on a separate page within a 40-sec. time limit. Following the completion of an unrelated 6-min. interpolated task which was utilized to control the time interval between original drawings and recall, each *S* was asked to draw as many of the designs as he could remember. All recalled designs were scored as correct or incorrect on the basis of a previously developed objective scoring system (4). A reliability check of 20 randomly selected protocols revealed 91.3% agreement in scoring by two judges.

Results and Discussion

Each *S* received three scores, the number of easy tasks (2, A, 1) correctly recalled; the number of medium difficulty tasks (6, 8, 5) recalled; and the number of hard tasks (3, 7, 4) recalled. The means and *SD*'s of these three scores for the High- and Low-Anxious groups are presented in Table 1. The results, expressed as percentages of correctly recalled tasks, are also presented in Fig. 1. These data were analyzed by a design (7, Type III) in which Anxiety and Sex were the between-*S*s factors and Task Difficulty was the within-*S*s factor. Table 2 summarizes the analysis of variance.

TABLE 1

Recall Scores of Hard, Medium, and Easy Tasks by Sex and Anxiety Level

(*N* = 44 in each group)

Task Diff.	High-Anxious *S*s		Low-Anxious *S*s	
	Mean	*SD*	Mean	*SD*
Hard	.82	.75	1.39	.91
Medium	1.82	.91	2.02	.81
Easy	2.41	.72	2.09	.90

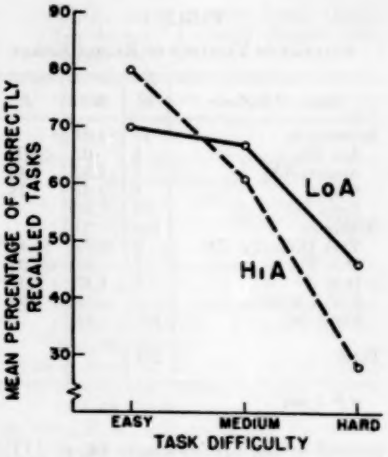

Fig. 1. The mean percentage of easy, medium, and hard tasks recalled by high- and low-anxious *S*s (*N* = 44 in each group).

As may be seen in Table 2, no significant effect was found attributable to Sex, in accordance with the findings reported in the earlier study (4). Therefore, for subsequent analyses and discussion, the data for the two sexes were combined.

The *F* for Task Difficulty was highly significant ($P < .001$) indicating that the designs as presented varied significantly in difficulty of recall. These results which reflect the combined effects of *both* design difficulty and serial position are in striking agreement with those re-

TABLE 2

ANALYSIS OF VARIANCE OF RECALL SCORES

Source of Variance	*df*	*MS*	*F*
Between *Ss*	87	1.07	
Sex (S)	1	.02	<1.00
Anxiety (A)	1	1.52	1.43
S × A	1	3.41	3.25
Error (b)	84	1.05	
Within *Ss*	176	.93	
Task Difficulty (D)	2	30.73	56.49**
D × S	2	.35	<1.00
D × A	2	4.37	8.03**
S × A × D	2	.56	1.06
Error (w)	168	.54	
Total	263		

** $P < .001$.

ported in the earlier study (**4**, p. 233). For all *Ss* combined it was found that the easy tasks were recalled 75% of the time, the medium tasks 64% of the time, while the hard tasks were recalled only 36% of the time.

The obtained *F* for the interaction of Task Difficulty and Anxiety was also highly significant ($P < .001$). An inspection of the means in Table 1 and Fig. 1 indicates that the LoA *Ss* performed relatively better on the hard tasks than the HiA *Ss*, while the HiA *Ss* were superior on the easy tasks. A *t* test of the mean differences indicates that the differences between the HiA and LoA *Ss* on the hard tasks was significant ($t = 3.16$, $P < .01$) and the difference between these groups on the easy tasks approached significance ($t = 1.88$, $P < .10$). The difference on the medium difficulty tasks was not significant.

It should be noted that the measures in several of the treatment categories evidenced varying degrees of deviation from normality and moderate degrees of heterogeneity of variance. However, as Lindquist (**7**) notes, even extreme violations of the assumptions of normality and homogeneity of variance do not have an appreciable effect on the distribution of the *F* ratio. Thus, it is safe to conclude that the significant *F* values reported in Table 2 reflect true treatment and interactive effects.

The present results offer strong confirming evidence concerning the interactive relationship between drive and performance. With tasks of increasing complexity, presumably involving more or stronger competing responses, the performance of anxious (high drive) *Ss* should be relatively more impaired than that of nonanxious (low drive) *Ss*. These results are regarded as consistent with the findings of several recent studies from the Iowa laboratories (**3, 9, 12, 16**) demonstrating essentially the same phenomenon.

Since some recent studies have reported negative relationships between anxiety and measures of intelligence (e.g., **5, 8, 13**) and since recall performance has been shown to be a function of intelligence (**10**), it should be noted that there was no significant difference in the total recall of the two anxiety groups. The mean number of designs correctly recalled by the HiA and LoA *Ss* was 5.05 and 5.50, respectively, with the difference between these means nonsignificant (evaluated as the main effect of Anxiety, $F = 1.44$, $P > .20$). There is no basis at all for assuming an interactive relationship between intelligence and performance on recall tasks such as used in the present study that would account for our results.

As a further check on the effect of intelligence on these results, the ACE Psychological Examination results that were available for 21 HiA *Ss* and 22 LoA *Ss* were compared. The mean ACE score for the HiA group was 124 while the mean for the LoA group was 119; the difference between these means was not significant ($t = .72$), nor was the correlation between the ACE scores and the total recall scores for these 43 *Ss* ($r = .09$). Thus, in the absence of a main effect of anxiety or any relationship between intelligence and anxiety or intelligence and total recall, it must be concluded that the obtained difference in performance cannot be accounted for in terms of the usual alternative hypothesis of original differences in

intelligence, memory, or ability to learn. These findings are clearly consistent with deductions from Hullian theory (1, 2, 6, 11) based upon a drive interpretation of anxiety scale scores.

Summary

On the basis of Hullian theory and previous empirical research, an interactive relationship was hypothesized between Manifest Anxiety Scale scores and performance on the recall of a previously graded series of nine visual-motor tasks when no instructions to learn were given. The *Ss* (44 men and 44 women), who had extreme scores on the Taylor MAS, were given the nine designs of the Bender-Gestalt Test in a specified order which maximized the range of difficulty of recall of these tasks. Following an unrelated interpolated task, *Ss* were asked to redraw the designs they could remember.

The results show a significant interaction between MAS score and task difficulty; the high-anxiety *Ss* performed better on the easy tasks while on the more difficult tasks the low-anxiety *Ss* were better. These results were interpreted as supporting the hypothesized interactive relationship between task difficulty and drive as measured by MAS scores.

References

1. Farber, I. E. Anxiety as a drive state. In M. R. Jones (Ed.) *Nebraska symposium on motivation*. Lincoln, Nebr.: Nebraska Univer. Press, 1954.
2. Farber, I. E. The role of motivation in verbal learning and performance. *Psychol. Bull.*, 1955, **52**, 311–327.
3. Farber, I. E., & Spence, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 120–125.
4. Goodstein, L. D., Spielberger, C. D., Williams, J. E., & Dahlstrom, W. G. The effects of serial position and design difficulty on recall of the Bender-Gestalt Test designs. *J. consult. Psychol.*, 1955, **19**, 230–234.
5. Grice, G. R. Discrimination reaction time as a function of anxiety and intelligence. *J. abnorm. soc. Psychol.*, 1955, **50**, 71–74.
6. Hull, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
7. Lindquist, E. F. *Design and analysis of experiments*. Boston: Houghton Mifflin, 1953.
8. Matarazzo, J. D., Ulett, G. A., Guze, S. B., & Saslow, G. The relationship between anxiety level and several measures of intelligence. *J. consult. Psychol.*, 1954, **18**, 201–205.
9. Montague, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45**, 91–96.
10. Peek, R. M., & Olson, G. W. The Bender-Gestalt recall as an index of intellectual functioning. *J. clin. Psychol.*, 1955, **2**, 185–188.
11. Spence, K. W. A theory of emotionally based drive (*D*) and its relation to performance in simple learning situations. *Amer. Psychologist*, 1958, **13**, 131–141.
12. Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 116–119.
13. Spielberger, C. D. On the relationship between manifest anxiety and intelligence. *J. consult. Psychol.*, 1958, **22**, 220–224.
14. Taylor, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
15. Taylor, J. A., & Chapman, J. P. Anxiety and the learning of paired associates. *Amer. J. Psychol.*, 1955, **68**, 671.
16. Taylor, J. A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **44**, 61–64.

(Received September 3, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

ASSOCIATIVE CONNECTION BETWEEN PAIRED VERBAL ITEMS

JACK RICHARDSON AND ADRIENNE ERLEBACHER

Harpur College *Northwestern University*

The meaningfulness or association value of single words (5), nonsense syllables (1, 3, 4), and consonant syllables (8) is specified by the frequency of association to each item separately but only adjectives have been rated for similarity of meaning and associative connection between pairs of items. The high positive correlation found by Haagen (2) indicates that instruction to rate pairs of adjectives for similarity of meaning and instruction to rate pairs of adjectives for associative connection produce approximately the same ranking of the pairs.

The difficulty of learning paired-associate verbal lists increases as the meaningfulness decreases. However, with the materials available it is not possible to vary meaningfulness and associative connection in order to determine the relationship between the two. It is plausible that associative connection tends to increase as the meaningfulness of the items of the pair increases. This study rates the associative connection between pairs of items selected from different levels of meaningfulness.

PROCEDURE

Subjects and instructions.—The *Ss* were students in three introductory psychology courses at Northwestern University during the summer of 1956. The booklets containing the pairs of items were distributed so that approximately half of the *Ss* in each class rated the pairs according to one of the two sets of instructions. The rating was done during the class period. Each booklet consisted of nine pages with 24 pairs listed on the left margin of each page. Each pair was followed by the numbers 1 through 15 which were used as the scale for rating.

The first page of each booklet consisted of instructions to *S*. Two different sets of instructions were used but each *S* rated the pairs by only one of the two sets. Both instructions included a list of 12 pairs of items at the top of the page consisting of four pairs each of adjectives, nonsense syllables, and consonant syllables. This was followed by an example of the 15-point scale and an explanation of how it was to be used. One set of instructions specified that the rating was to be according to meaningfulness, namely, "You are to rate each pair as to how meaningful it would be to you if it were inserted into the list above . . . . Meaningfulness may be defined as the degree to which the items of a pair denote the same or similar objects, actions, or conditions." The other set of instructions were in terms of how fast *S* thought he could learn the pairs. "You are to rate each pair according to how fast you think you could learn it if it were inserted into the list above . . . . If you were given a pair to learn, you would study the pair until, when presented with the first member of the pair alone (bal), you could give the second member of the pair (vic), just as you would if you were learning a foreign language vacabulary." The instructions were followed by six practice pairs which were rated before *S* opened the booklet.

Selection of items.—The three types of items were selected from three different sources: the adjectives from the Thorndike-Lorge word list (7), the nonsense syllables from Glaze (1), and the consonant syllables from Witmer (8). The two items of a pair were always of the same type and of either high or low values of frequency or association value. Thus the pairs were either adjectives, nonsense syllables, or consonant syllables and either high-high (H-H), low-high (L-H), high-low (H-L), or low-low (L-L) with the first value designating the meaningfulness of the first or stimulus item and the second designating the meaningfulness of the response item when conceived of as paired associates.

The pairs of items were selected in blocks of 12 by the following procedure and this was repeated 18 times for a total of 216 pairs. The types of pairs (adjectives, nonsense syllables, consonant syllables) were randomly assigned to sequence within the block of 12 pairs subject

TABLE 1

MEDIANS OF MEAN RATINGS OF PAIRS OF ITEMS UNDER LEARNING INSTRUCTIONS AND MEANING INSTRUCTIONS

Meaningfulness of Items	Adjectives		Nonsense Syllables		Consonant Syllables	
	L Inst.	M Inst.	L Inst.	M Inst.	L Inst.	M Inst.
H-H	13.29	11.07	7.76	5.10	3.20	2.76
L-H	11.93	9.40	6.20	3.84	2.94	2.38
H-L	12.02	9.32	6.12	4.06	2.94	2.38
L-L	10.93	7.95	5.62	3.25	2.59	2.11

to the restrictions that there could not be three of a type in a sequence and that each type must appear four times within the block. The meaningfulness pairings (H-H, L-H, H-L, L-L) were randomly assigned subject to the restrictions that there could not be three of a kind in the sequence and that no two identical combinations (type of item and meaningfulness pairing) appear in the block. The first letter of each item was always a consonant and was randomly assigned subject to the restrictions that no letter was used more than twice within the block of 12 pairs and that letters were not repeated in adjacent pairs. Since there are few adjectives beginning with *q*, *x*, *y*, or *z*, these four letters were treated as one in the assignment of first letters, and items beginning with either of the four letters were acceptable.

The items were then selected from the appropriate source keeping the high values as high as possible and the low values as low as possible. If more than one adjective corresponded to the specifications, then the shortest was always used. The frequency of the adjectives, according to Thorndike-Lorge, was not less than 35 per million for the high frequency adjectives and not more than 15 per four million for the low frequency. The high association value nonsense syllables had a value of at least 80% and the low a maximum of 27%. The corresponding values for the consonant syllables were 71% and 21%.

RESULTS

The medians of the mean ratings for the pairs in each category are given in Table 1. It is apparent that the ratings for adjectives, nonsense syllables, and consonant syllables differ for both sets of instructions. In fact there is practically no overlap of the distribution of the three types of pairs except for the nonsense syllables and consonant syllables rated for meaningfulness. The L-L pairs are consistently rated lower than the H-H pairs with the L-H and H-L pairs falling between the two. There is little consistency in the small differences between the L-H and H-L pairs. The Friedman two-way analysis of variance (**6**, p. 166) of the ranks of the medians in Table 1 gave χ^2_r of 8.2 for both sets of instructions. The probability is .017 that the medians of the various combinations of high and low values of stimulus and response were drawn from the same population.

The median rating obtained by Learning instructions is 6.33 while that for the Meaningfulness instructions is 3.96. A median test gives a χ^2 of 30.17 indicating that the Learning instructions result in consistently higher ratings. The product-moment correlation coefficient between the mean ratings of each item by the two sets of instructions is .96 and the r within each type of material is .82, .81, and .59 for the adjectives, nonsense syllables, and consonant syllables, respectively.

The 223 pairs of items are presented in Table 2 and the meaningfulness of each stimulus and response presented separately. *AA* indicates that the word occurs more than 100 times per million words, *A* that the word appears between 50 and 100 times per million, and the numerical entries following the words indicate first the frequency and then the unit in millions. Thus 18/9 indicates that the word appears 18 times in nine million words (**7**). The association value of the nonsense syllables

TABLE 2

MEAN RATINGS OF 223 PAIRS OF ITEMS UNDER LEARNING INSTRUCTIONS AND MEANING INSTRUCTIONS

No.	Pair*		Frequency or Association Value		Instructions					
					Learning			Meaning		
	Stimulus	Response	S	R	*N*	Mean	*SD*	*N*	Mean	*SD*
1	first	new	AA	AA	56	13.86	2.57	54	13.30	3.52
2	valuable	real	A	AA	56	13.82	2.49	53	12.91	3.72
3	fine	warm	AA	AA	56	13.48	2.98	54	12.65	3.55
4	direct	proper	AA	AA	56	13.25	3.08	54	12.43	3.87
5	ten	young	AA	AA	55	13.04	3.08	54	12.09	4.08
6	safe	green	AA	AA	55	13.33	3.36	53	12.04	4.17
7	heavy	chief	AA	AA	56	13.78	2.32	53	11.85	4.18
8	pleasant	few	AA	AA	55	13.13	3.20	54	11.44	4.01
9	many	keen	AA	35/1	56	12.73	3.03	54	11.31	4.18
10	great	due	AA	AA	56	13.11	2.73	54	10.83	4.28
11	vast	less	A	AA	55	13.16	3.39	53	10.81	4.66
12	merry	hot	A	AA	55	13.82	2.35	54	10.61	4.42
13	happy	late	AA	AA	55	13.51	2.71	54	10.48	4.67
14	well	red	AA	AA	55	13.98	2.73	54	10.43	4.71
15	wild	soft	AA	AA	56	13.14	3.13	53	10.26	4.83
16	major	various	A	AA	55	12.38	3.36	54	10.20	4.97
17	just	cold	AA	AA	56	13.66	2.96	54	10.11	4.71
18	bad	six	AA	AA	56	13.12	3.07	54	10.11	5.04
19	kinky	queer	5/4	48/1	55	12.82	3.11	53	11.09	4.36
20	jellied	white	10/4	AA	56	12.61	3.15	53	10.60	4.37
21	diabetic	more	4/18	AA	56	12.16	2.96	53	10.55	4.40
22	tarsal	hard	4/18	AA	56	12.11	3.49	53	10.15	4.68
23	goodish	deep	4/18	AA	55	12.25	2.84	54	10.04	4.63
24	finny	back	4/18	AA	56	12.68	3.16	54	9.87	4.82
25	visceral	fit	4/18	AA	27	11.59	3.17	28	9.64	4.61
26	styptic	yellow	4/18	AA	56	12.20	3.28	54	9.61	4.81
27	basaltic	sick	4/18	AA	54	11.76	3.90	54	9.57	5.08
28	tangy	waste	4/18	AA	29	12.62	2.88	26	9.23	5.44
29	wabbly	big	9/4	AA	55	11.98	3.59	54	9.13	5.01
30	fusty	wise	4/18	AA	55	11.87	3.30	54	8.63	4.78
31	nascent	dry	12/4	AA	29	11.62	3.87	26	8.35	4.99
32	ciliary	poor	4/18	AA	56	11.20	4.05	54	8.35	4.85
33	lorn	top	9/4	AA	27	11.70	3.30	28	8.14	4.63
34	rabbinical	pretty	4/18	AA	55	10.69	4.08	53	8.02	4.85
35	yeastly	glad	4/18	AA	55	11.62	3.42	54	7.56	4.72
36	besprent	daily	4/18	AA	56	10.75	3.50	53	6.79	4.94
37	long	towery	AA	4/18	55	12.93	3.24	53	12.66	3.40
38	human	recuperative	AA	4/18	56	12.02	3.53	54	12.20	3.99
39	dark	nutty	AA	14/4	56	13.12	2.62	54	11.35	4.56
40	final	wifely	A	11/4	56	12.07	3.33	52	10.65	4.77
41	past	zestful	AA	4/18	55	12.38	3.45	54	10.44	4.58
42	cool	ghoulish	AA	4/18	56	11.71	3.87	53	10.23	4.45
43	base	theatric	AA	4/18	56	12.36	2.94	53	10.19	4.66
44	plain	largo	AA	4/18	56	12.96	2.72	53	10.17	4.56
45	close	tenpenny	AA	4/18	56	12.18	3.41	54	9.52	4.84
46	wide	pursy	AA	4/18	56	11.82	4.01	54	9.33	4.72
47	waste	tangy	AA	4/18	27	12.15	2.99	28	9.32	4.73
48	side	kinetic	AA	9/4	55	11.71	3.24	54	9.28	4.36
49	bay	hipped	AA	4/18	55	12.38	3.97	53	9.02	4.89

* The frequency of the words and the association value of the syllables are given separately for the stimulus and response items. The number of *Ss* and the *SD* of the ratings are given separately for the two sets of instructions.

TABLE 2—*Continued*

No.	Pair*		Frequency or Association Value		Instructions					
					Learning			Meaning		
	Stimulus	Response	S	R	*N*	Mean	*SD*	*N*	Mean	*SD*
50	dry	nascent	AA	12/4	26	11.23	3.38	28	8.68	4.11
51	next	loamy	AA	4/18	56	11.50	3.71	53	8.47	4.79
52	vain	semiarid	A	4/18	55	11.44	3.48	54	8.37	4.90
53	kind	volant	AA	4/18	55	12.38	3.27	54	8.33	5.33
54	mad	residuary	A	4/18	56	10.59	3.99	52	8.27	4.99
55	fit	visceral	AA	4/18	29	11.62	3.47	26	8.23	5.08
56	gray	pivotal	AA	4/18	56	10.49	4.30	53	8.08	4.83
57	top	lorn	AA	9/4	29	11.14	3.80	26	7.62	4.92
58	wiggly	tunable	4/18	4/18	55	12.13	3.21	54	10.07	4.92
59	nudist	waxy	6/4	9/4	55	12.67	3.07	54	9.76	4.91
60	loutish	tricksy	4/18	4/18	55	10.93	3.76	54	9.50	4.88
61	sneaky	diluvial	4/18	4/18	56	11.05	3.49	53	9.02	4.94
62	jocose	newsy	8/4	4/18	56	10.20	3.94	54	8.46	4.89
63	measly	carking	4/18	4/18	55	11.33	3.44	54	8.44	4.69
64	tonal	wobbly	4/18	6/4	55	10.87	3.70	53	8.30	5.05
65	nosy	vulpine	4/18	4/18	56	11.30	3.83	53	8.13	5.04
66	knowable	lateen	6/4	9/4	56	10.98	3.56	53	7.77	5.25
67	whilom	ritualistic	6/4	4/18	55	9.09	4.44	54	7.69	4.74
68	citrous	vibratory	4/18	4/18	56	11.07	3.29	54	7.56	4.79
69	refrangible	wigged	4/18	6/4	56	9.95	3.95	54	7.48	5.20
70	trine	jittery	4/18	5/4	55	10.93	3.47	54	7.06	5.22
71	xanthous	jumbo	4/18	7/4	56	8.66	4.36	54	6.85	5.02
72	kaleidoscopic	pluvial	14/4	4/18	56	8.59	5.00	53	6.81	5.39
73	wonted	neap	15/4	6/4	55	10.51	3.67	53	6.45	4.86
74	fem	hos	93	93	55	8.22	4.56	53	6.47	4.82
75	tel	cum	93	93	55	8.89	4.68	54	6.39	5.22
76	wiz	sec	87	93	55	6.99	4.35	53	5.94	4.73
77	ruf	mus	100	93	56	7.16	4.46	54	5.57	4.67
78	qiz	sem	87	100	56	8.07	4.77	54	5.26	4.99
79	fal	tex	93	100	56	9.50	4.12	54	5.24	4.07
80	vol	det	93	93	56	7.96	4.18	54	5.20	4.62
81	dep	ber	100	100	56	8.45	4.31	52	5.19	4.29
82	bec	sev	100	93	56	7.68	4.20	54	5.13	4.09
83	sic	jil	93	87	56	8.23	4.46	54	5.06	3.89
84	ken	jel	100	100	56	7.96	4.47	53	4.98	3.77
85	pil	mon	100	100	56	7.84	3.95	53	4.85	3.79
86	fev	mir	93	93	56	6.48	3.87	54	4.41	3.71
87	rus	kip	100	80	55	6.82	4.22	54	4.39	3.99
88	pav	kof	93	80	54	7.59	4.56	54	4.28	3.89
89	tep	laz	87	93	55	7.44	4.53	54	4.06	3.74
90	lan	kuc	93	87	56	6.80	4.06	54	3.85	3.48
91	bel	vif	100	100	56	6.99	4.08	53	3.75	3.36
92	bep	lis	13	100	55	7.16	4.25	54	4.54	4.18
93	poh	sav	20	93	29	7.03	4.36	26	4.42	3.86
94	gic	wip	13	87	56	7.11	4.67	53	4.25	3.72
95	saj	vic	20	100	56	6.27	4.38	53	4.25	3.73
96	wuc	hov	20	87	56	6.45	4.27	54	4.24	3.72
97	ruy	bis	20	93	55	6.44	4.20	53	4.19	3.89
98	jec	hin	13	93	56	6.64	3.83	53	4.09	3.90
99	cux	kel	13	87	55	7.22	4.20	54	4.07	3.54
100	paf	cin	13	100	55	7.04	4.23	54	3.96	3.53
101	gih	pac	0	100	29	6.17	4.12	26	3.88	3.54
102	nuk	hur	20	100	27	5.92	4.44	28	3.79	3.20

* The frequency of the words and the association value of the syllables are given separately for the stimulus and response items. The number of *Ss* and the *SD* of the ratings are given separately for the two sets of instructions.

TABLE 2—*Continued*

No.	Pair[a]		Frequency or Association Value		Instructions					
					Learning			Meaning		
	Stimulus	Response	S	R	*N*	Mean	*SD*	*N*	Mean	*SD*
103	daq	cov	7	93	56	5.77	3.87	53	3.77	3.36
104	tef	biz	20	100	56	5.68	3.92	53	3.64	3.61
105	cej	gul	7	100	56	6.39	4.20	54	3.57	3.09
106	vak	jen	20	87	55	6.47	3.91	53	3.51	3.27
107	hif	coz	20	93	55	5.65	3.71	54	3.41	2.90
108	cij	dul	0	100	56	5.07	3.81	54	3.39	3.05
109	bej	wev	20	100	56	5.98	4.06	54	3.24	2.84
110	xej	fon	0	93	55	4.33	3.27	53	3.11	3.09
111	bip	qes	20	87	27	5.81	3.99	28	3.07	3.02
112	hod	bof	87	7	56	6.75	4.19	54	4.96	4.06
113	nav	ruq	100	13	55	6.45	3.91	53	4.75	4.45
114	bes	ceh	93	13	56	5.91	3.86	53	4.43	3.72
115	ger	juq	87	7	56	6.78	4.06	54	4.24	3.46
116	reg	kih	100	13	55	5.59	3.97	53	4.17	3.71
117	sav	poh	93	20	27	7.07	3.78	28	4.14	3.57
118	fiv	mez	93	20	56	6.70	4.14	54	4.13	3.84
119	rud	gik	100	13	55	6.33	3.96	53	4.13	3.70
120	jav	heg	87	20	56	6.89	4.25	53	4.09	3.49
121	lak	ged	93	13	55	6.44	4.12	53	4.08	3.61
122	lec	meq	93	0	56	5.77	3.73	53	4.04	3.72
123	med	qug	93	0	56	5.07	3.77	54	3.98	3.51
124	vin	dej	100	20	55	5.38	3.81	54	3.89	3.79
125	hur	nuk	100	20	29	5.90	3.53	26	3.85	3.42
126	fav	piy	93	13	56	5.66	3.74	51	3.84	3.54
127	yeg	mec	80	7	54	5.63	3.60	54	3.70	3.15
128	ref	siw	93	20	55	6.62	4.16	54	3.44	2.92
129	qes	bip	87	20	29	6.34	4.41	26	3.27	2.96
130	tes	jiq	100	7	54	5.26	3.89	54	3.19	2.90
131	pac	qih	100	0	27	5.74	3.81	28	2.96	2.71
132	gov	nub	0	13	56	6.68	4.18	54	4.24	3.98
133	jat	leq	20	20	56	5.98	3.97	53	3.91	3.69
134	vaf	qap	0	0	55	5.87	4.07	53	3.85	3.71
135	kex	beh	13	20	56	6.53	3.93	54	3.59	3.13
136	guk	huq	7	20	55	5.96	3.93	53	3.55	3.62
137	woj	leb	7	20	55	5.69	3.96	53	3.53	2.85
138	paq	nij	27	20	56	5.25	3.71	54	3.44	2.99
139	peh	ruw	27	20	55	5.84	4.41	54	3.39	3.42
140	buq	diw	20	20	56	5.86	3.73	54	3.26	2.78
141	laj	vux	0	7	56	4.77	3.82	54	3.24	2.98
142	vej	wup	13	13	55	4.42	3.56	54	3.22	2.78
143	kuq	xat	13	0	55	4.91	3.74	54	3.09	3.30
144	gix	piw	7	7	56	5.54	3.92	54	3.06	3.03
145	rix	juc	13	13	55	5.78	3.81	54	2.98	2.68
146	wut	xip	20	0	56	5.32	4.08	53	2.92	3.01
147	dax	fep	0	20	56	5.48	3.82	52	2.71	2.61
148	tev	xef	13	0	56	5.02	3.70	53	2.68	2.35
149	xeq	gid	0	7	55	4.44	3.55	54	2.61	2.99
150	knd	tgr	88	83	55	4.38	3.67	53	4.09	4.47
151	qwt	hrd	71	100	56	3.48	3.59	53	3.36	3.74
152	crk	rmp	83	83	55	3.64	3.09	53	3.34	3.61
153	kng	nsh	92	88	56	4.80	3.90	54	3.27	3.34
154	dft	wsh	83	83	56	3.37	3.42	53	3.17	3.41
155	bgr	flx	83	83	56	3.20	3.00	54	3.02	3.26

[a] The frequency of the words and the association value of the syllables are given separately for the stimulus and response items. The number of *Ss* and the *SD* of the ratings are given separately for the two sets of instructions.

TABLE 2—*Continued*

No.	Pair[a]		Frequency or Association Value		Instructions					
					Learning			Meaning		
	Stimulus	Response	S	R	*N*	Mean	*SD*	*N*	Mean	*SD*
156	hnk	gbr	79	79	56	2.62	2.70	54	2.94	3.22
157	kgs	trp	71	92	55	3.13	2.55	54	2.83	3.12
158	xmp	dbt	67	92	56	2.57	2.71	54	2.76	2.94
159	chr	xct	83	71	56	2.89	2.69	53	2.68	3.05
160	whp	znc	100	88	55	3.82	3.52	54	2.67	3.08
161	zbr	dmp	79	88	56	3.02	3.00	54	2.61	2.76
162	czr	znk	79	71	55	2.76	2.57	54	2.59	2.92
163	qtz	blw	71	83	55	2.85	2.57	54	2.56	2.88
164	jng	ngd	79	79	56	2.80	2.94	54	2.41	2.10
165	txg	chs	83	79	56	3.30	2.76	54	2.37	2.70
166	rpm	gbs	79	79	56	3.23	3.23	53	2.13	2.11
167	xzf	jrn	0	83	29	2.65	2.94	26	3.12	3.64
168	cgj	xms	13	79	56	3.46	3.72	54	3.04	3.74
169	cxq	frd	8	92	55	2.84	2.73	53	3.00	3.58
170	fhc	mrk	13	83	56	3.45	2.98	54	2.91	3.06
171	sbh	brk	17	83	29	3.21	2.91	26	2.73	2.94
172	fjh	rng	8	92	56	3.57	2.66	54	2.63	2.47
173	hfk	pkg	8	92	56	3.84	3.66	52	2.50	3.10
174	nqx	fnd	17	92	55	2.98	2.99	54	2.46	2.97
175	hjx	phd	13	88	56	2.46	2.76	54	2.41	2.45
176	sxh	xlt	17	71	56	2.46	2.53	53	2.38	3.04
177	mjf	ngt	8	79	55	3.14	2.85	54	2.37	2.93
178	cjz	sch	17	83	27	2.37	1.64	28	2.36	2.98
179	sjh	rfd	17	92	56	3.14	2.81	53	2.32	2.64
180	zjf	hmp	0	83	55	3.14	2.81	54	2.26	2.13
181	fjs	pkr	17	79	56	3.18	2.73	53	2.26	2.50
182	wcq	brn	13	92	56	2.77	2.81	53	2.08	2.20
183	jqz	xtr	13	79	56	2.28	2.34	54	2.11	2.45
184	tcz	glr	17	83	55	2.56	2.61	54	1.96	2.32
185	gjx	qsh	4	71	56	2.77	2.81	52	1.88	2.03
186	mqj	bnk	13	92	27	2.89	2.48	28	1.79	1.63
187	bnk	mqj	92	13	29	2.90	2.48	26	3.62	3.25
188	drk	hjc	100	8	55	3.44	2.91	52	3.25	3.82
189	slp	fpw	92	13	55	3.05	3.10	53	2.89	3.27
190	sch	cjz	83	17	29	3.07	3.05	26	2.65	2.72
191	sld	tzx	88	17	55	3.07	2.81	54	2.59	3.01
192	hrt	cfj	92	13	56	2.89	2.57	54	2.56	2.64
193	grl	nqs	92	17	56	3.34	3.16	53	2.57	2.77
194	fwn	sbq	83	13	56	3.25	2.99	51	2.57	2.48
195	grp	xbq	88	4	55	2.47	2.64	53	2.43	2.86
196	jnt	pxz	83	21	56	2.59	2.62	54	2.39	2.63
197	qlt	wqh	83	13	55	3.36	3.35	53	2.38	2.64
198	jlt	hbm	83	17	55	3.49	2.92	54	2.31	2.46
199	wng	bqf	92	13	56	2.87	2.93	54	2.30	2.33
200	brk	sbh	83	17	27	2.92	2.69	28	2.21	1.95
201	grt	sgj	92	17	56	2.96	2.89	54	2.19	2.49
202	trn	xjq	88	0	56	2.80	2.70	53	2.08	2.58
203	nbr	kqx	79	4	56	2.54	2.69	54	2.06	2.31
204	jrn	xzf	83	0	27	3.07	2.51	28	2.00	1.85
205	fsk	cjx	96	8	56	2.93	2.85	53	1.98	2.20
206	fth	szj	83	13	55	2.33	2.59	54	1.85	2.11
207	hfx	cgq	42	13	55	2.65	2.58	53	2.58	3.33
208	fjw	sjz	17	21	56	2.53	2.51	54	2.46	2.86

[a] The frequency of the words and the association value of the syllables are given separately for the stimulus and response items. The number of *Ss* and the *SD* of the ratings are given separately for the two sets of instructions.

TABLE 2—*Continued*

No.	Pair[a]		Frequency or Association Value		Instructions					
					Learning			Meaning		
	Stimulus	Response	S	R	*N*	Mean	*SD*	*N*	Mean	*SD*
209	hkm	wbq	13	8	56	2.61	2.40	54	2.41	2.66
210	cqh	mhj	8	8	56	3.02	2.84	53	2.38	2.62
211	fcj	rjx	8	17	55	2.45	2.61	54	2.31	2.67
212	nlq	gqb	21	8	56	2.80	2.75	53	2.28	2.41
213	szk	bqd	21	21	55	2.85	3.60	54	2.17	2.63
214	tzh	bhj	8	8	55	2.44	2.54	53	2.17	2.56
215	mhz	khx	4	4	56	2.71	3.06	54	2.11	2.11
216	mcf	qxn	13	4	56	2.59	2.60	54	2.11	2.31
217	zxj	qxb	0	4	56	2.16	2.37	54	2.11	2.26
218	pjx	sfj	21	21	55	2.47	2.50	54	1.93	2.01
219	mwb	zqj	8	0	56	2.20	2.37	52	1.92	2.21
220	jxb	cqz	13	17	56	2.18	2.38	54	1.89	2.18
221	pzw	bjq	17	17	54	2.70	2.80	53	1.83	1.82
222	mbx	zfq	17	4	56	2.49	2.67	52	1.81	1.89
223	bqj	ncf	4	17	56	2.93	2.63	54	1.61	1.47

[a] The frequency of the words and the association value of the syllables are given separately for the stimulus and response items. The number of *S*s and the *SD* of the ratings are given separately for the two sets of instructions.

(1) and consonant syllables (8) are given for each item. The number of *S*s who rated each pair, the mean rating, and the *SD* are presented separately for the learning instructions and the meaning instructions. A small *N* indicates that the same pair appears in the table with the stimulus item and response item reversed.

The ratings obtained indicate that associative connection is not independent of meaningfulness. The higher the meaningfulness of the pairs the higher the associative connection. The studies concerned with the effect of meaningfulness on learning usually compare the learning of two different types of material or of items with different association values. It is possible that the facilitation of learning by meaningfulness is not the result of association value per se but of the increase in associative connection resulting from the increase in meaningfulness. The important thing may not be the number of associations which can be given to the stimulus and response items independently but the relationship between the two.

Summary

Pairs of words, nonsense syllables, and consonant syllables of different degrees of meaningfulness were rated under two different sets of instructions. The meaning instructions specified that the rating was to be on the degree of similarity of meaning of the two items in a pair while the learning instructions specified rating on the basis of how fast *S* thought he could learn the pair.

Learning instructions gave a higher rating than meaning instructions but the correlation between the ratings obtained under the two sets of instructions was .96. The ratings varied with the meaningfulness of the items and with the type of material. The 223 pairs of items and their ratings are presented.

References

1. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
2. Haagen, C. H. Synonymity, vividness, familiarity and association value ratings

of 400 pairs of common adjectives. *J. Psychol.*, 1949, **27**, 453–463.

3. Hull, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, **45**, 730–734.
4. Krueger, W. C. F. The relative difficulty of nonsense syllables. *J. exp. Psychol.*, 1934, **17**, 145–153.
5. Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.
6. Siegel, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.
7. Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words.* New York: Teachers Coll., Columbia Univer., 1944.
8. Witmer, L. R. The association value of three-place consonant syllables. *J. genet. Psychol.*, 1935, **47**, 337–360.

(Received September 20, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

CHANGES IN DISTRIBUTION OF MUSCULAR TENSION DURING PSYCHOMOTOR PERFORMANCE [1]

LEE W. GREGG

Carnegie Institute of Technology

The purpose of the present research was to determine whether or not changes in generalized tension accompany changes in proficiency of psychomotor task performance. It was supposed that measures of generalized tension could be derived from separate recordings of muscle-action potentials obtained from a number of muscle groups, some of them being quite remote from the locus of the activity required for the performance of the task.

A number of experiments (**1, 3, 4, 6, 9, 12**) indicate that specific patterns of muscular activity are established during the performance of motor tasks. Freeman (**3**) observed a progressive isolation of muscular activity in the body member most directly involved in a pursuit task with practice on the task. This result implies a decrease in tension in body members less directly involved, a finding that was confirmed in a study by Henderson (**6**). A different kind of result has been obtained in several seemingly diverse experiments such as the studies of Davis (**2**), Malmo, Shagass, and Davis (**7**), Meyer (**8**), Smith (**10**), and Stroud (**11**). In each of these experiments, changes in muscular tension were found to be related to degree of learning or to variables suggesting changes in the motivation of *S*. It therefore seems reasonable to hypothesize that changes in generalized tension will accompany changes in proficiency of performance.

[1] This research was carried out under Contract DA19-129-QM-250 between the Quartermaster Research and Development Command and Carnegie Institute of Technology. The writer is indebted to W. H. Teichner and E. R. Dusek of the contracting agency for their guidance and suggestions. Graduate assistants who participated in these studies were E. Prien, E. Holshouser, I. Goldberg, and A. Chenzoff.

METHOD

Experimental Design

Two separate experiments were carried out to investigate the relationships between changes in the distribution of tension and changes in proficiency of performance. In both experiments an arm-hand steadiness task and a tapping-aiming task were used to evaluate task differences.

Experiment I.—In Exp. I, 12 *Ss* received practice on both of the tasks for five consecutive days. The first day was a practice day in which *Ss* received instructions for the tasks and were familiarized with the method of measuring the muscle-action potentials. On Days 1 and 2, the first and second experimental sessions, *Ss* received three practice trials of 1-min. duration on each of the tasks. The order in which they performed the tasks was counterbalanced so that half of the *Ss* practiced tapping first followed by the steadiness test on Day 1 and the reverse order of tasks on Day 2. The remaining *Ss* received the opposite order of tasks. Days 3 and 4, the third and fourth experimental sessions, were a repetition of the first two sessions except that a packboard load of 67 lb. was attached to *Ss*. Muscle-action potentials were recorded from each of five separate body locations during the performance of the task and in rest periods preceding and following the performance. The experiment was designed in such a way that task effects, practice effects for days and trials, and the effect of the pack load could be assessed.

Experiment II.—Experiment II attempted to produce decrement in the proficiency of performance on the psychomotor tasks by requiring *Ss* to carry out work that would fatigue specific muscle groups. Six groups of 10 *Ss* each were used. Three groups received practice on the tapping task only while the other three groups received practice on the steadiness task. A

control condition of rest was interpolated between two performance periods for one of the groups from each task. The conditions of interpolated work for the other groups consisted of squeezing a hand dynamometer or lifting a 4-kg. weight with the arm extended. Thus, the design was a 2×3 factorial with 10 *Ss* randomly assigned to each of the six treatment combinations. Task differences at two levels and interpolated activity at three levels comprised the main independent variables. Muscle-action potentials were recorded from four separate muscle groups during each of the six performance trials of the single experimental session. In addition, potentials were recorded at the beginning and at the end of the session while *Ss* were resting.

Apparatus

The electronic system that was used to measure and record the muscle-action potentials (MAP's) consisted of a high-gain biological amplifier described previously by Gilmer, Caywood, and Shaw (5), a switching arrangement that permitted successive samples from five separate body locations, and five electronic integrators from which quantitative readings in meter microamperes corresponding to the total magnitude of potential change in the separate body locations could be obtained. In addition, an electronic timing circuit was employed that permitted any number of successive sweeps of the switching device. This timing circuit was adjusted, for both experiments, to give four sweeps in 1 min. and provided audible clicks that signaled the beginning and end of each practice trial for the psychomotor tasks.

A calibration run on the integration was carried out before every experimental session. During the 1-min. period in which the electrodes were open-circuited, noise in the amplifier and switching circuits contributed 8 micro-amp. ± 2 to the meter values. The values obtained in the calibration runs prior to each session were subtracted from the values obtained during each session to give an adjusted score.

In order to obtain performance measures for the steadiness and tapping tasks, an electronic chronoscope was constructed. Four Berkeley decade counting units, Model 705A, counted the positive peaks of a 100-cps signal from a Hewlett-Packard Oscillator calibrated against the line frequency. Control of the chronoscope was accomplished through a silent relay in the plate circuit of a paralleled 6SN7 twin triode. Whenever *S* touched his stylus to the plates of the tapping task or to the sides of the aperture in the steadiness task, the fixed bias in the triode grids was shorted and the chronoscope activated. Thus, error scores in .01 sec. were obtained without any appreciable current or voltage change in the shielded experimental room.

The psychomotor tasks, tapping and steadiness, were the same for both experiments. The tapping task required *S* to move the stylus back and forth over a 1-in. hurdle placing the tip of the stylus into the $\frac{1}{4}$-in. holes in the centers of 4-in. square plates. The center-to-center separation of the holes was $4\frac{1}{2}$ in., since the width of the hurdle was $\frac{1}{2}$ in. between the plates. Rate of tapping was determined by the beat of a metronome set at 100 taps/min. The arm-hand steadiness task required *S* to hold the $\frac{1}{4}$-in. diameter stylus within the $\frac{1}{4}$-in. diameter aperture on a plate mounted in an upright rectangular box. In both of the experiments and for both of these tasks, a trial consisted of a 1-min. sample of performance during which MAP's were also obtained.

Surface electrodes were used to pick up the potential changes. Five pairs of bipolar electrodes constructed of silver plate $\frac{1}{2}$ in. in diameter were attached to the various parts of the body by means of Ace elastic bandages. The skin area was prepared by washing with alcohol and then applying standard electrode jelly to a carefully delimited part lying just below the electrode. The spacing between electrodes was approximately $1\frac{1}{2}$ in. In Exp. I, where repeated measures were required on successive days, five electrode locations were marked with a violet dye: the left biceps (M_1); the trapezius laterally to the right side of the neck (M_2); the frontalis just above the right eye (M_3); the right biceps (M_4); and the right forearm flexors (M_5). The same muscle groups were used in Exp. II with the exception of the frontalis group (M_3).

Procedure

Experiment I.—The first day was a practice day. The *Ss* were shown the apparatus for measuring the MAP's and assured that the device was harmless. The electrodes were attached to *S* who was seated at the table in the electrically shielded room. Instructions for carrying out the tapping and steadiness tasks were given. A mark on the table just to the left of the task devices positioned *S*'s left hand and arm. The right hand and arm were placed in a similar position on the right. The *S* was then told to sit and relax until *E* told him to "pick up the stylus." This was to be followed by the signals "Ready . . . Go!" At the end of each practice trial, the signal "Stop!" was given and *S* was to put down the stylus and return the right hand to the rest position. Three

practice trials on each of the tasks were given each *S* on the practice day. On the other days, the following sequence was carried out. Three successive 1-min. samples of MAP's were obtained, designated first rest period (R_1). A second rest period (R_2) followed immediately during which three more 1-min. samples were obtained. As far as *S* was concerned, these first two conditions were as one long relaxation period of approximately 12 min. For one-half of the *S*s in the first session, three 1-min. performance trials on the steadiness task (T_1) followed the initial rest periods. During this time MAP's and the error score in .01 min. were recorded. These *S*s then received three 1-min. trials on the tapping task (T_2). The other half of the *S*s received the practice trials on the tasks in the reverse order. The first experimental session terminated with a rest period (R_3) in which three 1-min. samples of the MAP's were obtained. On Day 2 these conditions were repeated except that the order of the tasks was reversed for all *S*s, thus counterbalancing tasks over the two sessions. Days 3 and 4 were an exact replication of Days 1 and 2 except that after the initial rest (R_1) the pack board was attached and remained so until the end of each session.

TABLE 1

ANALYSIS OF VARIANCE OF THE INTEGRATED ACTION POTENTIALS, EXP. I

Source	*df*	*MS*	Error	*F*
1. Muscles (M)	4	29,663.54	35	11.16*
2. Trials (T)	2	21.27	32	.68
3. Conditions (C)	4	48,304.62	36	23.03*
4. Days (D)	3	746.83	20	1.84
5. M×T	8	27.25	32	.87
6. M×C	16	6,843.13	33	43.23*
7. M×D	12	45.48	33	.29
8. T×C	8	22.72	32	.72
9. T×D	6	29.31	32	.93
10. C×D	12	127.68	33	.81
11. M×T×C	32	15.30	32	.49
12. M×T×D	24	20.32	32	.96
13. M×C×D	48	75.00	32	2.39*
14. T×C×D	24	19.57	32	.62
15. M×T×C×D	96	6.37	31	.55
16. Subjects (*S*)	11	2,552.09	34	3.88*
17. S×M	44	1,135.82	33	7.18*
18. S×T	22	23.28	32	.74
19. S×C	44	371.16	33	2.35*
20. S×D	33	405.58	33	2.56*
21. S×M×T	88	14.42	32	.46
22. S×M×C	176	222.05	32	7.06*
23. S×M×D	132	192.27	32	6.12*
24. S×T×C	88	15.75	32	.50
25. S×T×D	66	12.41	32	.39
26. S×C×D	132	69.57	32	2.21*
27. S×M×T×C	352	8.18	31	.71
28. S×M×T×D	264	14.29	31	1.24*
29. S×M×C×D	538	46.42	31	4.01*
30. S×C×T×D	264	18.02	31	1.55*
31. S×M×T×C×D	1056	11.57		
Total	3599			
32. Pooled 28,29,30	1066	31.43		
33. Pooled 13,22,23,26	488	158.29		
34. Pooled 17,19,20	121	658.60		
35. Pooled 6 and 17	60	2,657.77		
36. Pooled 6 and 19	60	2,097.02		

Note.—The significance of all terms is tested against the next higher order term that is significant and that contains that term. Where more than one higher order terms are significant, a pooled error term is used.
* $P = .025$.

Experiment II.—Only a single experimental session was required for each *S*. After a few preliminary remarks concerning the nature of the MAP apparatus, each *S* was seated at the table in the shielded room and the electrodes were attached. The *S*s were given instructions for the particular psychomotor task assigned to them and the system of signals used in Exp. I. In the initial rest condition (R_1) three 1-min. samples of the MAP's were obtained while *S*s sat relaxed. This was followed by a series of three 1-min. practice trials on the task assigned (T_1) during which MAP's were obtained and the error scores recorded. Next, the interpolated activity was introduced. There were three interpolated activities: 5-min. rest (A_0 and B_0); squeezing a hand dynamometer (A_1 and B_1); and lifting a 4-kg. weight (A_2 and B_2). The steadiness task is A and the tapping task is B. The *S*s were requested to squeeze the dynamometer to 20 kg. as many times as possible or to lift the weight by flexing the extended right arm at the elbow as many times as possible. They were then asked to try again. The number of squeezes or lifts was recorded. After this work, three more practice trials were given on the assigned task (T_2). A rest period (R_3) concluded the session.

The *S*s for these experiments were male undergraduate students at Carnegie Institute of Technology. The *S*s in Exp. I were paid $1.00 per hr. The *S*s of Exp. II were "volunteers" from the ROTC program.

RESULTS

Experiment I

Muscle-action potentials.—An analysis of variance was carried out using the integrated MAP's obtained during each 1-min. trial of the experimental session. The raw data consisted of the 3600 separate measures of integrated potentials adjusted by subtracting from each meter reading a value obtained in a calibration run preceding the experimental session. These scores were entered into a five-way classification table and an analysis in the form of a five-factor

TABLE 2

MEAN INTEGRATED ACTION POTENTIAL MEASURES FOR EXP. I

	Muscle Groups						Muscle Groups						Mean 1 & 2
	M_1	M_2	M_3	M_4	M_5	Mean	M_1	M_2	M_3	M_4	M_5	Mean	
	Day 1						*Day 2*						
R_1	11.5	25.0	21.5	9.6	10.6	15.6	10.0	26.7	21.6	8.9	10.0	15.4	15.5
R_2	11.8	26.5	21.6	9.0	10.5	15.9	10.5	27.2	21.2	8.3	9.2	15.3	15.6
T_1	18.5	36.2	24.8	36.3	30.3	29.2	16.3	37.6	23.3	36.5	38.5	30.4	29.8
T_2	16.2	38.5	25.4	35.6	36.5	30.4	17.7	38.7	23.2	39.4	38.8	31.6	31.0
R_3	9.9	25.5	22.0	9.3	9.9	15.3	9.5	28.0	22.0	8.5	9.7	15.5	15.4
Mean	13.6	30.3	23.1	20.0	19.6	21.3	12.8	31.6	22.3	20.3	21.2	21.6	21.5
	Day 3						*Day 4*						
R_1	13.0	26.1	22.9	8.8	11.4	16.4	10.8	26.6	23.6	7.9	10.2	15.8	16.1
R_2	12.4	29.9	24.1	9.6	12.2	17.7	11.5	29.9	24.1	8.3	12.2	17.2	17.4
T_1	18.1	39.1	24.7	41.3	37.7	32.2	19.9	38.2	26.3	40.6	43.1	33.4	32.8
T_2	16.7	36.0	24.0	39.0	36.3	30.4	18.2	38.0	23.8	42.0	37.9	32.1	31.2
R_3	12.5	28.4	24.1	9.0	12.3	17.3	11.2	30.4	25.4	10.3	11.5	17.7	17.5
Mean	14.5	31.9	24.0	21.5	22.0	22.8	14.1	32.6	24.8	21.8	23.0	23.3	23.0

factorial design was performed. The *Ss* were treated as one classification. In the summary table of this analysis, Table 1, Muscles represent the electrode locations for the five muscle groups investigated. Trials are the three successive measures obtained within each condition of the experimental session. These Conditions were the initial rest (R_1), a second rest (R_2), performance on the arm-hand steadiness test (T_1), performance on the tapping-aiming test (T_2), and a final rest (R_3). Days represent the four experimental sessions. The *Ss* wore the pack load on Days 3 and 4 except during the initial rest (R_1).

In Table 1, a number of *F*'s are significant at the .01 level. The quadruple interaction variances involving *Ss* provide error variances and are used to test the significance of other sources. The rationale underlying the pooling of error variances is that a source be tested against the next higher order term (or terms) that contains the source and is significant. Individual differences among *Ss* are reflected by the significant *F* for *Ss* (Line 16). The interactions of *S* with Muscle groups, Conditions, and Days (Lines 17, 19, 20, 22, and 23) were also significant sources of variation. Over-all differences in MAP's obtained from different muscle groups are shown by the significant *F* for Muscles and differences in the MAP levels attributable to the nature of the activity are shown by the significant *F* for Conditions. Of particular interest are the significant M × C and M × C × D interactions. These interactions indicate the differential effects of task performance on the magnitude of the integrated MAP's obtained from muscle groups involved directly in the task as contrasted with the the more remote, inactive muscle groups.

The nature of the significant effects

may be found by examination of Table 2, a table of the mean values of the integrated MAP's for each Muscle, Condition, and Day. For each day, mean values for the conditions obtained by summing over muscles are presented. On Day 1 the initial rest (R_1) yielded a mean of 15.6 units. R_2 on the same day gave a mean of 15.9 units, and R_3, at the end of experimental session, 15.3 units. Contrasted with these values are the means obtained during the performance trials on the arm-hand steadiness test (T_1) and the tapping task (T_2), which were 29.2 and 30.4 units, respectively. Differences in MAP's obtained from the different muscle groups are shown by the mean values obtained by summing over conditions. Actually, a more adequate description of the activity of the separate muscle groups can be found by inspection of the resting potentials only, since there is a differential change in the magnitude of the potentials from the muscle groups directly involved in the performance of the task from resting to working conditions. This is the significant M × C interaction effect. During rest periods only, the left biceps (M_1), right biceps (M_4), and right forearm flexors (M_8) show potentials of approximately the same level. For the initial rest, the mean values range from 7.9 to 13.0 units. Of greater magnitude were the potentials from the frontalis (M_3), the means ranging from 21.5 to 23.6 units for the initial rest only. The trapezius (M_2) yielded the largest mean MAP's ranging from 25.0 to 26.7 units.

TABLE 3

ANALYSIS OF VARIANCE OF THE INTEGRATED ACTION POTENTIALS, EXP. II

Source	*df*	*MS*	Error	*F*
1. Tasks (T)	1	170.08	4	.27
2. Work (W)	2	772.17	4	1.23
3. T × W	2	7.29	4	.01
4. *Ss* within (*S*)	54	628.07	20	3.23*
5. Conditions (C)	3	33,106.36	7	8.61*
6. Muscles (M)	3	25,741.74	7	6.70*
7. C × M	9	3,842.81	16	3.39*
8. T × C	3	118.07	16	.10
9. T × M	3	885.13	16	.78
10. W × C	6	147.93	16	.13
11. W × M	6	77.01	16	.07
12. T × C × M	9	158.25	16	.14
13. T × W × C	6	7.33	16	.00
14. T × W × M	6	104.50	16	.09
15. W × C × M	18	91.66	16	.08
16. T × W × C × M	18	1,134.50	20	5.84*
17. *S* × C	162	88.56	19	2.94*
18. *S* × M	162	299.76	19	9.95*
19. *S* × C × M	486	30.12		
Total	959			
20. Pooled 17, 18	324	194.17		

Note.—The significance of all terms is tested against the next higher order term that is significant and that contains that term. Where more than one higher order terms are significant, a pooled error term is used.
* $P = .025$.

Performance measures.—Separate analyses of variance were carried out for the tapping and steadiness tasks. The median performance scores for the three trials within each session were entered into a Day by *Ss* classification table. For the steadiness test, Days ($F = 5.99$, 3 and 33 *df*) and *Ss* ($F = 60.07$, 11 and 33 *df*) were significant beyond the .01 level. The means computed from the median scores for each day were 7.62, 5.26, 7.47, and 6.16 sec. for the four sessions. For the tapping task, Days ($F = 10.56$, 3 and 33 *df*) and *Ss* ($F = 19.93$, 3 and 33 *df*) were again significant beyond the .01 level. These means were 6.67, 5.34, 4.64, and 4.13 for the four sessions. For both tasks, some learning is evident. The elevated scores for the steadiness test on Days 3 and 4 suggest that the packboard load adversely affected performance for this task. However, in the case of the tapping task, the design is inadequate for separation of the effects of load and practice.

Experiment II

Muscle-action potentials.—In Exp. II the Tasks provided the treatment

TABLE 4

MEAN INTEGRATED ACTION POTENTIAL MEASURES FOR EXP. II

	Rest					Hand Dynamometer					Lifted Weight				
	M_1	M_3	M_3	M_4	Mean	M_1	M_3	M_3	M_4	Mean	M_1	M_2	M_3	M_4	Mean
Arm-Hand Steadiness (Task A)															
R_1	8.7	22.9	16.4	12.8	15.2	11.4	25.1	13.6	15.2	16.3	10.6	20.9	24.5	10.7	16.7
T_1	10.8	39.7	51.9	25.7	32.0	15.6	40.6	60.5	31.8	37.1	10.5	38.0	50.2	30.9	32.4
T_2	11.5	43.9	52.7	27.7	34.0	20.4	43.7	61.3	29.0	38.6	11.3	51.0	51.1	37.5	37.7
R_2	7.3	24.9	18.7	8.0	14.7	8.8	24.4	12.3	16.5	15.5	8.6	22.6	20.6	7.8	14.9
Mean	9.6	32.8	34.9	18.6	24.0	14.0	33.4	36.9	23.1	26.9	10.2	33.1	36.6	21.7	25.4
Tapping-Aiming (Task B)															
R_1	10.2	19.8	14.7	11.9	14.2	9.8	25.6	14.1	15.2	16.2	9.7	22.4	12.0	13.3	14.4
T_1	11.7	34.7	48.5	37.8	33.2	12.7	42.1	49.9	49.4	38.5	12.0	36.5	50.0	35.6	33.5
T_2	10.5	35.2	47.8	35.8	32.3	15.7	41.7	49.7	40.5	36.9	13.1	44.6	50.6	34.2	35.6
R_2	7.6	22.2	12.9	8.1	12.7	9.1	24.4	10.6	11.4	13.9	8.7	25.5	11.2	9.5	13.7
Mean	10.0	28.0	31.0	23.4	23.1	11.8	33.4	31.1	29.1	26.4	10.9	32.2	31.0	23.2	24.3

categories for one variable and the kind of interpolated Work—rest, squeezing the hand dynamometer, or lifting the weight—provided the levels of the second variable. The mixed analysis of variance of the integrated MAP's also included the experimental Conditions of rest or performance and Muscle groups. These factors were completely orthogonal to Tasks and interpolated Work so that the final analysis yielded sources of variation as shown in Table 3.

The $S \times C \times M$ interaction was used as the basic error term to test the $S \times C$ and $S \times M$ interactions. Since these sources were significant, they were pooled (Line 20) to give an error term that was used to test the significance of the four-way interaction $T \times W \times C \times M$ and *Ss* within treatment combinations (Line 4). Both of these sources were significant and were therefore used to test certain effects as indicated in the summary table.

No significant variation in the integrated potentials for the main effects of Task and interpolated Work was obtained. Over-all differences in potentials from muscle group to muscle group and for the Conditions, rest periods vs. performance periods, were found. In addition, the significant $C \times M$ interaction indicates the differential amounts of change in the integrated potentials from rest to performance conditions for the different muscle groups. As in Exp. I, the trapezius shows the greatest resting potentials while the active muscle groups in the right arm give the highest potentials during the performance of the tasks. Inspection of Table 4, a table of means of the three-trial medians for the 10 *Ss* in each treatment group, suggests the nature of these differential effects.

Performance measures.—In order to determine the effect of the interpolated activity on the proficiency of performance for each of the psychomotor tasks, the difference between the median error scores for

the two performance periods—i.e., before and after the interpolated activity—was obtained. The differences were obtained by subtracting the three-trial median for each *S* in the first practice period (T_1) from the three-trial median obtained following the interpolated activity (T_2). A positive difference indicates a decrement in proficiency while a negative difference represents an improvement. For the steadiness task, mean values of −.52, 1.08, and 7.38 sec. were obtained for the rest, hand dynamometer, and lifted-weight conditions, respectively. For the tapping task, means of −.84, −.15, and .34 were found for those conditions in the same order.

An analysis of variance was carried out on these differences plus a constant value of 4 to eliminate negative numbers. For both tasks, the effect of interpolated activity was significant. For the steadiness task ($F = 66.72$, 2 and 27 *df*) a slight improvement was obtained from the first to the second practice periods with interpolated rest. A decrement was found after squeezing the hand dynamometer and a still larger decrement obtained after lifting the weight. For the tapping task ($F = 3.34$, 2 and 27 *df*) all *S*s showed an increase in proficiency from the first to the second practice periods (as opposed to 7 out of 10 *S*s improving on the steadiness task) with interpolated rest. Some general improvement was also observed after squeezing the hand dynamometer, but less than that for the interpolated rest. A decrement was found for the tapping task following the interpolated work of lifting the 4-kg. weight.

Discussion

During performance of the tasks all muscle groups, even those not engaged in overt task-related activity, exhibited higher potential values. A simple measure of generalized tension in terms of the mean value obtained by summing over the individual muscle groups seems appropriate for the subsequent discussion.

Task effects.—The steadiness task differed from the tapping task in that the former required a relatively static positioning response while the latter involved a repetitive movement pattern. Generalized tension as measured by the sum of the action potentials for the separate muscle groups was not different for these tasks. The relative amount of involvement of the muscle groups was also essentially the same for both tasks. That the patterns of tension were so similar suggests that the over-all postural adjustments were very much alike.

Load and fatigue effects.—In Exp. I introduction of the pack load tended to increase generalized tension as indicated by the higher mean values of the potentials on Days 3 and 4. This tendency was not statistically significant; however, the effect was observed for both resting and performance trials.

In Exp. II an attempt was made to induce performance decrement by fatiguing specific muscle groups. Measures of generalized tension for the performance trials before and after the interpolated work or rest periods revealed no significant changes.

Generalized tension as related to proficiency of performance.—The conditions of load and fatigue were set up on the assumption that proficiency in the performance of the psychomotor tasks would be altered. In Exp. I, the effect of the pack load was not the same for the two tasks. A decrement was observed for the steadiness task, but continued improvement was found for the tapping task. Increases in generalized tension for the steadiness task appear to accompany the performance decrement. Improvement in tapping may be accompanied by a decrease in generalized tension.

In Exp. II, the changes in performance for the steadiness task ranged from

a slight improvement with interpolated rest to a decrement of considerable magnitude as a result of lifting the 4-kg. weight. Slight, statistically insignificant increases in the generalized tension measure were found for all three groups. For the tapping test, improvement in performance was obtained for both the resting condition and the hand dynamometer. Decreased proficiency in tapping resulted only in the group that lifted the 4-kg. weight. Decreases in tension accompanied improvement in tapping and an increase in generalized tension occurred for the group in which a decrement was found.

Summary

Changes in the magnitude of integrated MAP's obtained from various muscle groups were observed in two separate experiments in which the experimental conditions were designed to produce differential proficiency in motor performance. An arm-hand steadiness task and a tapping-aiming task were used.

In Exp. I, 12 *Ss* practiced on the two tasks for four consecutive days. MAP's from five muscle groups were recorded. On Days 3 and 4, *Ss* wore a packboard load totaling 67 lb. Experiment II was designed to produce performance decrement by fatiguing the specific muscle groups most directly involved in the tasks. Ten *Ss* were assigned to each of six treatment groups. Interpolated between performance trials on each task was a period of work (or rest for the control group). The conditions of interpolated activity were: squeezing a hand dynamometer, designed to fatigue the right forearm flexors; and lifting a 4-kg. weight to fatigue the right biceps and upper arm.

The results were:

1. During performance of the motor tasks, an increase in the muscle-action potentials over the resting levels for all muscle groups was found. The amount of tension and its pattern over the various muscle groups was not essentially different for the two tasks.

2. Changes in this generalized tension during performance that could be attributed to the conditions of load or fatigue were not found.

3. The conditions of load and fatigue altered proficiency of performance of the tasks in such a way that increases in tension appeared to accompany performance decrement.

References

1. Daniel, R. S. The distribution of muscular action potentials during maze learning. *J. exp. Psychol.*, 1939, **24**, 621–629.
2. Davis, R. C. The relation of muscle action potentials to difficulty and frustration. *J. exp. Psychol.*, 1938, **23**, 141–158.
3. Freeman, G. L. Spread of neuro-muscular activity during mental work. *J. gen. Psychol.*, 1931, **5**, 479–494.
4. Freeman, G. L. The optimal muscular tensions for various performances. *Amer. J. Psychol.*, 1938, **51**, 146–151.
5. Gilmer, B. von H., Caywood, W. P., Jr., & Shaw, V. G. A high-gain low-noise neuro-biological amplifier. *Amer. J. Psychol.*, 1949, **62**, 576–578.
6. Henderson, R. L. Remote action potentials at the moment of response in a simple reaction time situation. *J. exp. Psychol.*, 1952, **44**, 238–241.
7. Malmo, R. B., Shagass, C., & Davis, J. F. Electromyographic studies of muscular tension in psychiatric patients under stress. *J. clin. exp. Psychopath.*, 1951, **12**, 45–66.
8. Meyer, H. D. Reaction time as related to tensions in muscles not essential in the reaction. *J. exp. Psychol.*, 1949, **39**, 96–113.
9. Oldroyd, C. R., & Moskowitz, B. Optimal tensions for various types of task as a function of proximity to focal musculature. *Proc. Okla. Acad. Sci.*, 1952, **33**, 265.
10. Smith, A. A. An electromyographic study of tension in interrupted and completed tasks. *J. exp. Psychol.*, 1953, **46**, 32–36.
11. Stroud, J. B. The role of muscular tension in stylus maze learning. *J. exp. Psychol.*, 1931, **14**, 606–631.
12. Telford, C. W., & Swenson, W. J. Changes in muscular tension during learning. *J. exp. Psychol.*, 1942, **30**, 236–246.

(Received September 20, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

A RELATIONSHIP BETWEEN INCENTIVE MOTIVATION AND ABILITY LEVEL IN PSYCHOMOTOR PERFORMANCE [1]

EDWIN A. FLEISHMAN

Yale University

It is often assumed that certain kinds of motive-incentive conditions have a marked influence on learning and performance in a variety of human tasks. Certainly, attempts to manipulate the motivation level of individuals through special verbal instructions or inducements represent one of the most widely used techniques in controlling performance. McGeoch and Irion (5) have reviewed the experimental evidence on a variety of human tasks. There is some evidence that verbal manipulation of motive-incentive conditions may facilitate learning and performance in such tasks as addition, reading, and writing, for example.

There is little evidence, however, that supplementary motivation instructions are sufficient to modify proficiency in psychomotor performance. Noble (7), for example, found no effects of supplementary motivational instructions on performance in a continuous, two-hand tracking task. Bilodeau (1) studied skill acquisition on the Rudder Control Test (6), in which *S* makes compensatory pedal adjustments to keep a cockpit lined up with an indicated target. He found no effect of instructions emphasizing different forms of social competition.

The present study is a further attempt to manipulate incentive motivation in relation to psychomotor performance. The task used differed from those employed in previous studies. The most critical difference was thought to be the fact that the present task was a "self-paced" task. In the previous studies, *S*s' rate of response was determined by the pace set by the task (e.g., the speed of a moving target). It was thought that a task in which *S* had control of his own pace might be more susceptible to the experimental treatments.

Since negative results had been the rule, rather than the exception, it was decided to "load" the experimental instructions to contain a variety of motive-incentive statements. No attempt was made to evaluate the relative effects of the different kinds of incentives involved.

A further objective was to explore a possible interaction between the effects of motivation incentives of this type and level of ability on the task.

METHOD

The *S*s were 400 basic trainee airmen at Lackland Air Force Base. All *S*s received training on the Complex Coordination Test (6). Essentially, *S* must make appropriate control adjustments of stick and rudder controls in response to successively presented patterns of visual signals. The visual signals appear on the upright panel. The upper set of lights indicates the correct position of the stick control in the right and left dimension; the center, upright rows of lights indicate the correct position of the stick in the to-and-from dimension; the bottom row of lights indicates the

[1] This study was carried out while the writer was at the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. The work reported was done under ARDC Project No. 7707, Task No. 27051. Permission is granted for reproduction, translation, publication, use, and disposal in whole or in part by or for the U. S. Government.

proper position of the rudder pedals. The task is to match the position of a stimulus light in each of these three dimensions by coordinate movements of the stick and rudder controls. A correct response (movement of stick and rudder to proper positions) is not accomplished unless both the hands and feet have completed and maintained for .5 sec. the appropriate adjustments simultaneously, at which point a new pattern of lights to be matched is automatically presented. The apparatus was modified to allow for trials 1 min. in length, separated by 20-sec. rest intervals. Two devices were placed in the experimental room, one on each side, with *E* and the control console in between. Thus, two *S*s were tested simultaneously. None of the *S*s had prior experience with these devices.

All *S*s were given identical initial instructions on how to perform the task and performed for five 1-min. trials. At the end of this initial practice period, treatment of half the 400 *S*s differed from treatment of the other half.

Half the *S*s received special "motivation" instructions, while the other half received none. The *S*s were assigned randomly to one of the two groups, the only restriction being that the assignment was made in groups of two *S*s at a time.

The "unmotivated" group received a 5-min. rest period prior to Trial 6. The "motivated" group received the following instructions during this 5-min. period:

"Now we really want to see how well you can do. Remember, this test measures how well you can carry out instructions and coordinate your action and ideas.

"Therefore, the results of this test are very important to you. They have an important bearing on your future assignment in the Air Force. So you must do your best. Thus far you have not been doing as well as you should, so you must try harder. You can improve your score by concentrating on what you are doing. Everyone will be told later how he has come out in relation to the others who are taking the test. A good way to get a high score is to try and beat the other fellow. I will also tell each of you who did better as soon as you are finished. Remember, work as quickly as you can, trying to complete as many matchings as possible. You will be in here for quite a while, but you will probably enjoy the test. The sooner you improve to a certain score, the sooner you will be allowed to leave.

"Ready, when the lights come on work as quickly as you can. Go!"

The *S*s in both groups then continued to practice on the complex task through 20 1-min. trials (separated by 20-sec. rest pauses). The *S*s in the "motivated" group received additional exhortations during certain of the rest periods following certain trials according to the following schedule: "[Trial 6] Remember, this test may have an important bearing on your future assignment in the Air Force. [Trial 8] Keep trying to improve your scores. [Trial 10] You can do better than that, work as quickly as you can. [Trial 12] You're doing better, but you'll still have to improve some more. [Trial 14] Keep trying to beat the other fellow. [Trial 16] Just a few more trials to go, if you keep improving. [Trial 18] Don't get careless, remember these scores count. [Trial 19] One more trial to go, make this your best one. [Trial 20] That is the end of the test."

The "unmotivated" *S*s were told nothing during any of the intervals between trials.

The number of successful matchings achieved during each trial was recorded for each *S*. The total scores obtained by each *S* during the initial five trials were obtained. This five-trial total was considered an index of the "initial ability" of *S* on this particular task. The *S*s in the "unmotivated" and "motivated" groups were then further stratified into high- and low-ability groups based on this initial ability measure. Those above the median were considered in the "high" ability groups and those below the median were considered in the "low" ability groups. A check on the matching achieved indicated that the mean initial "ability" level is practically identical for the "motivated" ($M = 15.95$, $SD = 5.77$) and "unmotivated" ($M = 15.97$, $SD = 5.21$) groups. It was also found that the "high-ability motivated group" ($M = 20.56$, $SD = 3.55$) is matched with the "high-ability unmotivated group" ($M = 20.09$, $SD = 3.02$) and that the "low-ability motivated group" ($M = 11.34$, $SD = 3.37$) is matched with the "low-ability unmotivated group" ($M = 11.85$, $SD = 3.34$). The correlations between performance on the matching segment of practice and subsequent practice trials has been found to be in the .70–.85 range in previous unpublished and published studies (e.g., **2**) with this task. The intertrial reliability of the matching segment of practice is above .90 (**2**).

Results

Table 1 presents the mean performance scores achieved by our four groups during the last 15 trials of practice (after introduction of the experimental condition) on the Complex Coordination Task. Table 2 presents the analysis of variance of these results.

TABLE 1

PERFORMANCE SCORES OF MOTIVATED AND UNMOTIVATED GROUPS

(N = 100 each subgroup)

Ability Level	Motivated		Unmotivated		Total	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
High	95.24	14.18	89.15	13.67	92.19	14.26
Low	67.72	14.08	66.95	13.72	67.34	13.90
Total	81.48	19.72	78.05	17.63	79.76	18.67

TABLE 2

ANALYSIS OF VARIANCE OF SCORES ON TRIALS 6–20 ON THE TASK

Source	*df*	*MS*	*F*
Motivation	1	1176.5	6.02*
Ability	1	6180.2	31.61*
Interaction	1	707.6	3.62*
Within cells	396	195.5	
Total	399		

* $P < .01$.

It can be seen that the main effects, ability level and motivation level, as well as their interaction effect, are significant beyond the .01 level when tested against the within-cell variance. The interpretation of the significant interaction effect is that the "motivation" treatment produced a difference among the high-ability *Ss* ($t = 3.09$, $P < .01$), but did not produce a difference among the low-ability *Ss*. This can be seen from inspection of Table 1.

Figure 1 presents the performance curves for each of the four groups. It can be seen that for the low-ability *Ss*, the curves of the motivated and unmotivated *Ss* overlap throughout the training period. However, for the high-ability *Ss*, the curves remain separated throughout the practice period, with the motivated high-ability *Ss* slightly, but consistently,

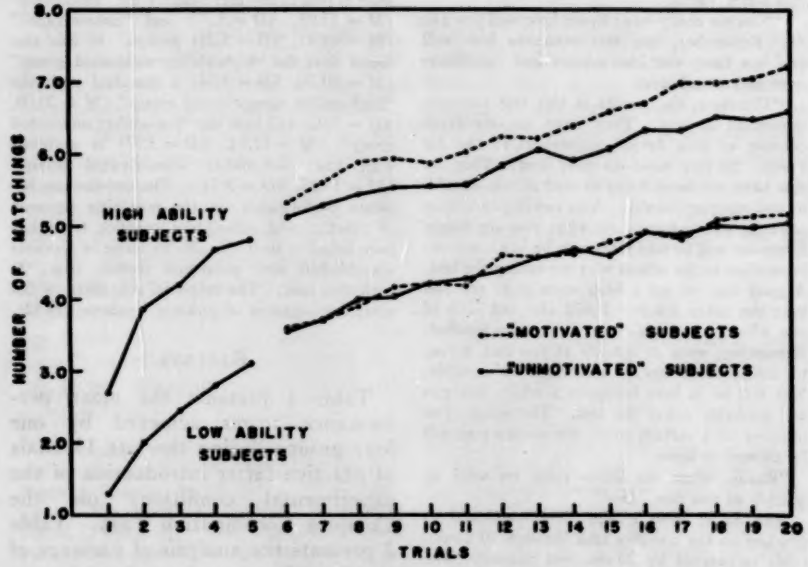

FIG. 1. Acquisition curves for subjects of different ability and motivation levels.

above the unmotivated high-ability *Ss*. While no statistical tests for trend differences were carried out, it is apparent from inspection that beyond Trial 6 the curves for each group are approximately linear and have similar slopes.

Discussion

The differences observed in the present study are not large, although significant, and should be interpreted with caution. This caution is also indicated by lack of corroborative evidence by other investigators with these kinds of tasks. However, French (4) has recent evidence of a significant interaction between ability level and an index of motivation in problem-solving performance. These results, together with ours, suggest the need for further experimentation into the generality of these findings.

It might be pointed out that the present study did not really compare "motivated" *Ss* with "unmotivated" *Ss*. There is a strong likelihood that basic trainee airmen serving as *Ss* in the laboratory are already highly motivated. The conditions of experimental testing are similar to those of classification testing, and as incoming recruits the *Ss* are extremely interested in the relationships of scores with future assignment. This initially high pre-experimental state of incentive motivation has been used as a possible explanation for the negative results by Noble (7) and Bilodeau (1). Thus, our treatments might better be described as a comparison of groups given "extra-motivation" instructions with groups already motivated, but not given these extra instructions. It is possible, of course, that the effects noted might be more marked if all *Ss* had not had this initial high level of motivation. Moreover, French (3) has shown recently that an individual's pre-experimental motivation level is a significant variable affecting the degree to which a desired level of motivation can be aroused by instructions within the experimental situation.

Summary

An attempt was made to produce differential effects of certain supplementary verbal motive-incentive instructions upon performance of individuals of different ability levels. The *Ss* ($N = 400$) practiced on a self-paced complex psychomotor task for 20 1-min. trials. The *Ss* were divided into four groups representing the possible combinations of levels of ability and motivation. Inspection of the learning curves indicates that each group improved at about the same rate and in an approximately linear fashion. However, the major finding was that although supplementary motivation instructions made no difference in the over-all performance of the low-ability *Ss*, these instructions did make a difference in the performance of the high-ability *Ss*.

References

1. Bilodeau, E. A. *Acquisition of skill on the Rudder Control Test with various forms of social competition.* San Antonio, Texas: Lackland Air Force Base, July 1951. (*USAF Hum. Resour. Res. Center Res. Note 51-6.*)
2. Fleishman, E. A., & Hempel, W. E. Changes in factor structure of a complex psychomotor test as a function of practice. *Psychometrika,* 1954, **19,** 239–252.
3. French, E. G. Some characteristics of achievement motivation. *J. exp. Psychol.,* 1955, **50,** 232–236.
4. French, E. G. Effects of interaction of achievement, motivation, and intelligence on problem solving success. *Amer. Psychologist,* 1957, **12,** 399–400. (Abstract)
5. McGeoch, J. A., & Irion, A. L. *The psychology of human learning.* New York: Longmans, Green, 1952.
6. Melton, A. W. (Ed.) *Apparatus tests.* Washington: U. S. Government Printing Office, 1947. (AAF Aviat. Psychol. Program Res. Rep. No. 4.)
7. Noble, C. E. An attempt to manipulate incentive-motivation in a continuous tracking task. *Percept. mot. Skills,* 1955, **5,** 65–69.

(Received September 23, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

INSTRUMENTAL AND COMPETING BEHAVIOR AS A FUNCTION OF TRIALS AND REWARD MAGNITUDE [1]

A. C. PEREBOOM AND B. M. CRAWFORD

Texas Technological College

In a study by Cotton (3) it was demonstrated that certain forms of competing or irrelevant behavior contribute more to the differences in total runway time as a function of hours of food deprivation than do changes in the vigor of the instrumental response itself—i.e., the speed of running. Is this a phenomenon peculiar to drive? Specifically, is it differences in the time involved in this instrumental response or in the time involved in all other runway behavior (retracing before reaching the goal, sitting, etc.) that produce the over-all changes due to (*a*) training and (*b*) different quantities of the incentive?

METHOD

Ten minutes of daily handling was begun approximately two weeks prior to training for two groups of albino rats (60–90 days old), with 10 in each group. They were also placed on a 23-hr. hunger drive at that time with 1 hr. a day of ad libitum Purina Chow feeding. Water was always available in the home cage.

Apparatus.—A 16-ft. open-top alley runway with 16-in. sides and 5 in. wide was painted a flat black throughout. It was illuminated by four shaded 15-w. bulbs mounted 30 in. above the runway floor with one each placed 6 in. and 5 ft. 6 in. from each end; no other lights were used. There were no goal box, entrance section, or nonretrace doors. Small porcelain cups fixed to 4-in. square plywood panels were placed at each end of this homogeneous alley.

Pretraining.—Adaptation occurred on one day and training the next. Adaptation consisted of three phases: runway exploration, prefeeding in the carrying cage, and prefeeding in the runway. The first phase began with a 10-min. period in the runway for pairs of *S*s followed by three 5-min. periods for individual *S*s, each period separated by 5–10 min. in the home cage. The *S* was placed directly in the center of the runway for the 10-min. period and the first 5-min. period, and was placed once at each end for the second and third 5-min. periods. After 10–20 min. of additional rest in the home cage, *S* was returned to the experimental room and given six feedings in his carrying cage from one of the runway cups. Each feeding here, in runway prefeeding and in training itself, consisted of either one 50-mg. pellet of Purina mash (.05-gm. group) or one 10-mg. pellet (.01-gm. group). A 50-sec. interval separated each feeding. Finally, after an additional 100-sec. rest in the carrying cage, *S* was given four feedings from each end of the runway in an alternating order: he was placed directly at the food cup (0 ft.) for the first four feedings, at 1 ft. from each end for the next two, and at 4 ft. from each end on the last two feedings. The *S* was removed from the runway 10 sec. after eating and was given a 50-sec. rest in the carrying cage between each feeding. His daily ration occurred 15–30 min. later in the home cage.

Training.—A shuttling procedure was used in training. The *S* was placed in the center of the runway and was permitted to obtain his first pellet from either end. A trial was defined as the time from the initiation of eating at one end of the runway to the occurrence of the same event at the opposite end. Thus, a "half-trial" exists at the beginning of training; it is not included in the present data. The *S* shuttled back and forth between the two cups for 40 trials without being touched by *E*.

An *E* sat at each end of the runway and placed a pellet in the food cup at one end when *S* began eating the pellet at the opposite end. Two electric timers, manually operated, recorded performance. One timer measured total runway time for each trial and the other measured cumulative forward progression time for each trial—i.e., the time when *S* was actually moving, however slowly, toward the appropriate food cup. Thus the difference between the total runway time for a given trial and the cumulative forward progression time for that same trial equals *S*'s competing response time.

[1] Research supported by a grant from the National Science Foundation (NSF-G3261).

Results

Median times as a function of training for the two components of runway behavior are shown in Fig. 1 for the two groups. In order to evaluate the incentive variable, the median competing time and the median instrumental time for the 40 trials was computed for each *S*.[2] The Mann-Whitney test gave two-tailed significance ($P < .001$) between the groups in the case of both response measures. The graphic method discussed by Moses (**10**, pp. 443–445) was then used in order to make statements about the magnitude of these differences and their variability. For competing behavior, the expected difference between the two groups was 19 sec. with a confidence interval ($P < .01$) of 7 sec. $\leqq \mu_{.01} - \mu_{.05} \leqq 55$ sec. For instrumental behavior it was 3 sec. and 1.5 sec. $\leqq \mu_{.01} - \mu_{.05} \leqq 5$ sec., respectively.

The Friedman two-way analysis of variance by ranks was used to test for the significance of trials. The data were the median times for blocks of eight trials for each *S* under each response measure. Although Fig. 1 indicates much more improvement under competing behavior than under instrumental behavior, all the acquisition curves are significant: instrumental times reached the same level of significance ($P < .02$) for both groups. Competing behavior was more significant for the .01-gm. group ($P < .001$) than for the .05-gm. group ($P \cong .01$).

The data appear to parallel Cotton's drive-variation results. The major component of runway behavior

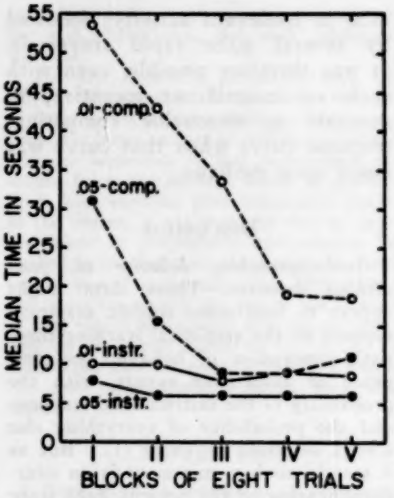

FIG. 1. Competing and instrumental acquisition curves for a group receiving .01-gm. rewards and a group receiving .05-gm. rewards.

that changes during acquisition or as a function of reward magnitude is the time involved in irrelevant or competing behavior, not the speed of the instrumental response itself. Even some of the differences in instrumental times that were obtained are artifactual.[3] The longer times frequently resulted when retracing occurred, with some of the forward progression time thus including time covering the same ground twice.

This tendency toward either performing within a restricted range of proficiency or indulging in considerable irrelevant activity was sometimes observed to occur over several trials as well as within a trial. Thus, some *S*s in the .01-gm. group frequently had inordinately long trials (10–25

[2] Actually, 23 *S*s were used in this experiment. Three *S*s were discarded when they appeared to be going to sleep in the runway. One was in the .05-gm. group and two were in the .01-gm. group.

[3] In fact, these "true" instrumental response times appear invariable enough that such responses may well approximate the maximum upper limit of performance in any given experimental situation.

min. in irrelevant activity) followed by several quite rapid traversals. It was therefore possible, even with such an insignificant incentive, to generate a reasonable competing-response curve when that curve was based upon medians.

Discussion

Goal-approaching behavior vs. goal-directed behavior.—These data might appear to lend some specific empirical support to the statistical learning theorist's conception of behavior as composed of all-or-none events, with the probability of the instrumental response and the probability of everything else always summing to unity (1). But as it stands, such a statement is an oversimplification of the present data since some variations in instrumental speeds do appear as a function of acquisition and reward magnitude. Observation also indicates such variations within a trial. Then what kind of significance should be placed on these small but real differences in instrumental speed? Part of them we have already noted as being due indirectly to retracing. It is now assumed that those remaining are due to slow, systematic, forward exploration which is confounded with the faster instrumental response itself (8). We are therefore assuming that these data, and Cotton's, have approached an ideal which may be difficult to realize; namely, that *S* does *not* select faster instrumental running responses with training, as implied by Logan (6), or with larger rewards, or stronger drives, but that the likelihood of such a response does increase.[4]

Training and reward magnitude as comparable variables.—Since both the reward magnitude variable and the acquisition variable produced similar results in this experiment, it should be possible to integrate their roles in the learning process. It is suggested that (*a*) the number of runway characteristics or cues that *can* evoke an instrumental running response is a function of the "amount" of runway (size and discriminable features) and how much experience *S* has had in that runway. (This will include both adaptation and training.) (*b*) The number of these cues that *will* evoke such a response when *S* is subsequently placed in that runway is a function of the amount of reward (size and discriminable features) and how much experience *S* has had with that reward. (This will include both prefeeding and training.) Some independent support for these parallel statements might be found in the fact that Hull was able to use the same function—albeit different parametric values—to relate reward magnitude to performance that he used to relate trials to performance (5).

We are suggesting that an associationistic or probabilistic interpretation be applied to learning about reward magnitude as well as learning about the runway. The emphasis therefore is upon the stimulus characteristics rather than the nutrient characteristics of the incentive (4): the larger the reward, the greater the potential number of goal stimuli to associate with runway stimuli ($_{s}H_{sG}$).[5]

Successive shifts in reward magnitude.—The preceding discussion has implications for performance as a function of successive amounts of reward in the same S-R situation. It was argued in an earlier paper that successive amounts of reward are successively learned (7). It was argued that to make such a

[4] This conclusion may have some relevance for human *S*s as well (9).

[5] In the case of Cotton's data, such remarks will need modification if we are to believe that drive is independent of learning (however, see 2): Suppose that the drive affects *S*'s reactivity to events, that it acts as a sensitizer (11). Then, with trials and reward magnitude constant, *S* may still learn no more under greater drive levels—i.e., he may still have the same number of runway stimuli-goal stimuli associations. Ideally then, he should run no faster; yet, as Cotton found, his probability of running would still be greater, more drive inducing a greater responsiveness to these anticipated goal stimuli when they are once aroused by their associated runway cues.

phenomenon possible *S* had to learn something distinctive about these different amounts of reward. That distinctiveness was postulated to reside in the "nature and/or intensity of the expectancy." But if this expectancy reduces to the number of goal stimuli associated with runway stimuli, at least with respect to trials and reward magnitude, how could a new and distinct second- or higher-order habit—following the discussion of that earlier paper—be created, be kept independent of a previous habit based on a different reward magnitude?

We can satisfy this "distinctiveness" requirement and thus tentatively solve our problem if we assume that in learning to perform in a given experimental situation for a given amount of reward, all runway cues associate with that reward (S_G's) *or* with something else in the goal area (S_O's). Then to *increase* the magnitude of reward during the learning process would change nothing in the case of those runway cues already associated with goal stimuli (s_1H_{SG}), but would suddenly introduce reward in the case of some of those cues now associated with goal area stimuli ($s_1H_{SO}H_{SG}$). To *decrease* the reward magnitude would change nothing in the case of those cues already associated with only goal area stimuli (s_1H_{SO}), but would introduce extinction for some of those cues now associated with the reward itself ($s_1H_{SG}H_{SO}$). It is interesting to note that the latter case maintains a natural continuum between reward reduction and reward removal.

Summary

A study of instrumental and noninstrumental acts in a straight runway was carried out using training and reward magnitude as the independent variables. Two groups of 10 rats each were given systematic adaptation on one day and 40 acquisition trials on the next. The *S*s were required to shuttle in a free-responding fashion between two food cups placed at opposite ends of a 16-ft. homogeneous alley runway. One group received a 50-mg. pellet for each traversal and the other a 10-mg. pellet. Instrumental (cumulative forward progression) and competing (retracing, sitting, etc.) times were recorded for each "trial"—i.e., for each reinforced traversal.

The data essentially parallel Cotton's (3) drive-variation results. Little selective learning of faster speeds occurred during acquisition or under the larger amount of reward, the probability of occurrence of that instrumental response being the primary source of change. Since both variables gave comparable results in this respect, it was suggested that an associationistic or probabilistic interpretation be applied to learning about the magnitude of the reward as well as learning about the runway: the larger the reward, the greater the number of goal stimuli to associate with runway cues. Implications of such a view for the learning of successively different amounts of reward within the same experimental situation were also discussed.

References

1. Bush, R. B., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.
2. Campbell, B. A., & Kraeling, D. Response strength as a function of drive level during training and extinction. *J. comp. physiol. Psychol.*, 1954, **47**, 101–103.
3. Cotton, J. W. Running time as a function of amount of food deprivation. *J. exp. Psychol.*, 1953, **46**, 188–198.
4. Guttmann, N. Operant conditioning, extinction, and periodic reinforcement in relation to concentration of sucrose used as reinforcing agent. *J. exp. Psychol.*, 1953, **46**, 213–224.
5. Hull, C. L. *Essentials of behavior.* New Haven: Yale Univer. Press, 1951.
6. Logan, F. A. A micromolar approach to behavior theory. *Psychol. Rev.*, 1956, **63**, 63–73.
7. Pereboom, A. C. An analysis and revision of Hull's theorem 30. *J. exp. Psychol.*, 1957, **53**, 234–238.
8. Pereboom, A. C. A note on the Crespi effect. *Psychol. Rev.*, 1957, **64**, 263–264.
9. Voeks, V. W. Acquisition of S-R connections: A test of Hull's and Guthrie's theories. *J. exp. Psychol.*, 1954, **47**, 137–147.
10. Walker, H. M., & Lev, J. *Statistical inference.* New York: Holt, 1953.
11. Wenger, M. A., Jones, F. N., & Jones, M. H. *Physiological psychology.* New York: Holt, 1956.

(Received September 30, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

STUDIES IN INCIDENTAL LEARNING: VII. EFFECTS OF FREQUENCY OF EXERCISE AND LENGTH OF LIST[1]

LEO POSTMAN AND PAULINE AUSTIN ADAMS[2]

University of California, Berkeley

Previous studies of serial learning under intentional and incidental conditions have shown that (*a*) incidental learning results in selective recall for a fraction of the items in the list, and (*b*) instructions to learn increase the strength of both correct and incorrect associations. Thus, intentional learners typically recall more than incidental learners but at the same time are more subject to intraserial interference (5). These findings suggest that two other variables which are known to influence serial learning—viz., frequency of exercise and length of list—should have differential effects on intentional and incidental learners.

Incidental *Ss* should benefit less from exercise than intentional *Ss*. Let us assume that amount retained is significantly influenced by *S*'s differential responses to the individual items in the series. The differential responses of incidental *Ss* tend to be reactions strongly associated with the learning items (6). Such responses are readily evoked and are likely to occur during the initial exposure to the series. Further presentations of the material will increase incidental learning to the extent that they (*a*) serve to strengthen initial differentiations of the items, and (*b*) provide opportunities for responses to additional items. These effects of exercise should, however, be limited. The initial differentiations presumably reflect strong habits which cannot benefit much from further exercise. The determinants of differential responses—e.g., the relative familiarity of the items—remain more or less constant from trial to trial, so that repetitions of the series should not add many new differentiations. By contrast, intentional *Ss*, who are instructed to learn the total series, cannot limit their differential responses to a few familiar items. They are also likely to verbalize their responses more explicitly and to engage in more rehearsal than incidental learners. Thus, successive trials should enable intentional *Ss* to increase the number of differentiations and connections among the items.

Earlier studies of the effects of exercise under intentional and incidental conditions have not produced conclusive results. Brown (1) tested intentional and incidental *Ss* by the anticipation method after 2 and 8 presentations of a verbal series and failed to find differential effects of frequency. His comparison was, however, between intentional *Ss* who had learned by the anticipation method from the beginning of practice and incidental *Ss* who were first introduced to the method at the time of the test. Saltzman and Atkinson (7) measured retention of a series of numbers by the method of recognition after frequencies of exposure ranging from 2 to 16 trials. The intentional *Ss* improved with practice whereas the incidental *Ss* did

[1] This research was facilitated by a grant from the Behavioral Sciences Division of the Ford Foundation. We wish to thank Estelle von Ende for her assistance.

[2] Now at Stanford University.

not. The differences between the two kinds of learners remained small until 16 presentations of the list had been given. The procedure of Saltzman and Atkinson was unfavorable to the detection of differences between intentional and incidental learners. Retention was tested by the method of recognition which is relatively insensitive to such differences (6). In addition, the orienting task performed during practice—finding and encircling each successive number in a coding matrix comprising the entire series—failed to control time of exposure to the materials and probably produced considerable intraserial interference, particularly under intentional conditions. Experiment I of the present study presents a further investigation of the effects of exercise under intentional and incidental conditions, with free recall as the measure of retention.

Consider next the effects of varying the amount of learning material. When retention is measured by the method of free recall, increases in the length of the list raise the absolute number, but at the same time lower the percentage, of items recalled (2, 8, 9). Thus, the relative efficiency of learning varies inversely with the length of the list. One explanation of this finding is that increases in the length of the list enhance the opportunities for intraserial interference. We have suggested that incidental learners are less subject to intraserial interference than intentional learners. It follows that increases in the amount of material should have less adverse effects on incidental than on intentional *Ss*. This advantage of the incidental learners should be more pronounced with nonsense lists than with meaningful lists since (*a*) nonsense series generate more intraserial interference than meaningful series and (*b*) when the materials are meaningful, incidental learning becomes less selective (6) and hence more susceptible to intraserial interference. These predictions are tested in Exp. II.

Experiment I

Method

Materials.—The learning materials consisted of a series of 24 nonsense syllables with association values of 26.67%–40.00% by Glaze's (3) norms. Twenty different random orders of the list were used. On a given trial, a different random order was used for each of the 20 *Ss* in a group. When there was more than one presentation, no order was used more than once per *S*.

Conditions of learning.—Different groups were exposed to the series 1, 2, 4, 8, or 16 times. There were an intentional group and an incidental group at each frequency of presentation. The experiment was introduced as a study of people's success in guessing numbers. The *Ss* were informed that a number between 1 and 10 had been assigned to each of the syllables, and that the assignment of these numbers had been determined by drawing cards from a well-shuffled deck. The groups receiving more than one presentation were also told that a new random assignment of numbers had been made for each trial. In a given trial, it was *S*'s task to guess which number had been assigned to each of the syllables.

Each *S* was handed a booklet with the number of pages corresponding to the number of presentations of the series. The syllables, numbered 1 through 24, were arranged in a column on each page. Each *S* also had a strip of cardboard with a window cut out of it, which he used to expose one syllable at a time in accordance with *E*'s instructions. When *E* called out the number "1," *S* slid the window down so as to expose the first syllable and wrote down the number which he guessed next to the syllable. This procedure was repeated for each successive syllable. The items were presented at a 4-sec. rate, with 6 sec. between trials.

Intentional *Ss* were informed that *E* was interested in studying the effects of memorizing on the accuracy of guessing. They were instructed to try to remember as many of the syllables as possible, without regard to the order of their appearance in the list. No learning instructions were given to the incidental *Ss*.

A 5-min. test of free recall was given 30 sec. after the end of practice. The *Ss* were requested

to write down as many of the syllables as they could remember, regardless of the order of presentation. They were urged to write down guesses as well as items about which they felt certain.

Subjects.—With five frequencies of presentation and two kinds of instructions, the experimental design included 10 groups of 20 *Ss* each. The *Ss* served in small groups, averaging four in number. They were assigned to the experimental conditions in rotation as they appeared in the laboratory, with correction for equal *Ns*. Twelve incidental *Ss* attempted to learn the syllables or expected a test of retention and were replaced. Of these, six were in the 16-presentation group, three in the 8-presentation group, two in the 4-presentation group, and one in the 2-presentation group. One incidental *S* did not follow instructions on the test of recall and was also replaced.

Results and Discussion

Amount retained.—The mean number of items correctly reproduced is shown in Fig. 1. The curve for intentional learners lies above that for incidental learners at all points. The intentional groups show steady improvement as a function of the number of presentations. The increases for the incidental groups are relatively small, and there is a slight decline between 8 and 16 presentations. Thus, intentional *Ss* benefit more from practice than incidental *Ss*, and the difference between the two kinds of learners increases progressively as a function of exercise.

The recall scores were subjected to analyses of variance after heterogeneity of variance had been removed by means of the Freeman-Tukey square-root transformation $(\sqrt{X} + \sqrt{X+1})$.[3] The differences among the intentional groups are clearly significant ($F = 26.71$, $df = 4$ and 95, $P < .001$). There is also a highly significant linear regression of retention scores on number of presentations ($F = 65.64$, $df = 1$ and 95, $P < .001$), indicating a reliable upward trend in the amount retained. The differences among the incidental groups fall just short of statistical significance. The F ratio is 2.42, whereas 2.46 is required for significance at the .05 level with 4 and 95 *df*. The variance associated with linear regression is not significant ($F = 1.45$). Exercise has only marginal effects on the amount retained by incidental learners.

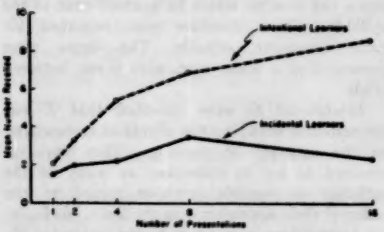

FIG. 1. Mean number recalled as a function of the number of presentations.

TABLE 1

OVER-ALL ANALYSIS OF VARIANCE OF RECALL SCORES IN EXP. I

Source	*df*	*MS*	*F*
Instructions	1	115.92	75.27*
Number of presentations	4	22.17	14.40*
Interaction	(4)	(8.75)	(5.68)*
Linear	1	30.35	19.71*
Deviation	3	1.55	1.01
Error	190	1.54	

* $P < .001$.

Table 1 presents an over-all analysis of variance of the retention scores. Both instructions and number of presentations are significant sources of variance, as is the interaction of these variables. The significant difference between the linear trends shows that the rate of improvement is reliably faster under intentional than under incidental conditions.

[3] This transformation was used prior to analysis of variance in all cases in which there was significant heterogeneity of variance.

TABLE 2

MEAN NUMBER OF DIFFERENT TYPES OF ERRORS AS A FUNCTION OF INSTRUCTIONS AND EXERCISE

Number of Presentations	Intralist Confusions				Other Errors				Total Errors			
	Intentional		Incidental		Intentional		Incidental		Intentional		Incidental	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	3.75	2.88	2.80	2.77	3.05	3.01	5.20	3.66	6.80	5.07	8.00	5.84
2	4.45	2.40	4.40	2.44	4.30	4.45	5.50	4.75	8.75	6.27	9.90	6.28
4	4.10	2.75	5.35	3.62	1.65	1.71	6.70	4.52	5.75	3.40	12.05	6.48
8	5.05	3.02	5.00	2.02	2.45	2.58	5.30	6.17	7.50	4.73	10.30	6.72
16	3.75	2.43	5.05	3.75	1.60	1.98	8.05	8.73	5.35	3.57	13.10	10.30

Frequency of errors.—Table 2 shows the mean number of wrong responses given by the various groups on the test of recall. At all levels of exercise, the frequency of errors is higher for incidental than intentional *S*s. Analysis of variance shows the over-all difference between the two conditions of learning to be significant ($F = 18.34, df = 1$ and 190, $P < .001$). The number of errors shows irregular variations as a function of exercise although the incidental groups give some evidence of an upward trend. Neither exercise nor the interaction of exercise and instructions is a significant source of variance ($F = .73$ and 1.82, respectively, $df = 4$ and 190 for both). Since incidental *S*s were not prepared for the test of recall, they probably attempted to maximize their scores by writing down a large number of guesses. This interpretation is supported by the fact that 11 incidental *S*s, and only two intentional *S*s, wrote down more than 24 syllables —i.e., more items than they knew to be in the list. This finding calls attention to the effect of an unexpected test of recall on the performance of incidental *S*s.

Types of errors.—The nature of the errors varies with the condition of learning. An attempt was made to distinguish between intralist confusions and other errors, with the latter probably including a large number of guesses. Two kinds of errors were classified and pooled as intralist confusions: (*a*) syllables in which two of the letters were correct and in their proper positions, but the third letter was incorrect and in a position appropriate to another syllable in the list, and (*b*) syllables in which all three letters were correct, but were reproduced in the reverse of the correct order. The mean numbers of intralist confusions and other errors are listed in Table 2. The relative frequency of intralist confusions is higher after intentional than incidental learning. Intralist confusions account for more than half of the errors of the intentional *S*s. After incidental learning, less than half the errors are intralist confusions. The relative frequency of confusion errors tends to increase with practice, particularly under intentional conditions. To determine the significance of these trends, the difference between the number of intralist confusions and other errors was determined for each *S*, and the difference scores were subjected to an analysis of variance. The two conditions of learning are significantly different ($F = 35.89$,

$df = 1$ and 190, $P < .001$). The effects of exercise approach significance ($F = 2.22$, whereas 2.41 is required for the .05 level of significance with 4 and 190 df). The interaction of instructions and exercise is not significant ($F = 1.09$). The distributions of errors are consistent with our assumption that intentional *Ss* are more subject to intralist interference than are incidental *Ss*. There are also indications that intraserial interference increases in the early stages of practice, just as it does in anticipation learning. It must be recognized, of course, that the validity of our criteria for identifying intralist confusions is unknown. It is reasonable to assume, however, that the frequency of the errors which were classified as confusions is directly related to the degree of intraserial interference.

Experiment II

Method

Materials.—The materials were lists of nonsense syllables and lists of two-syllable adjectives. The syllables had Glaze (3) association values between 26.67% and 40.00%. The adjectives were from Haagen's list (4) and fell within the highest interval on his 5-point scale of familiarity. For each type of material there were lists of 24, 36, and 48 items. The 24-item list of nonsense syllables was identical with that used in Exp. I.

Conditions of learning.—The orienting task consisted of matching each item to a geometric design. The *Ss'* task was to select one of the following geometric designs as best "fitting" a given item: square, triangle, circle, cross, arrow, diamond, and wavy line (sine curve). Large-scale drawings of these designs were in full sight throughout the experiment. The *Ss* were urged to make their judgments quickly and without attempting to make fanciful connections between items and designs. This task was chosen with a view to bringing out differential susceptibility to intralist interference as a function of length of list. Pairing of the different items in a list with the same set of geometric designs should favor response-produced generalization.

Each item was exposed for 4 sec. on a 2 × 2-in. slide. The height of the letters on the screen was 1.25 in. Following exposure of the slide, *Ss* had 6 sec. in which to draw the design they had chosen in the appropriately numbered space of a record blank. Thus, a new item was presented every 10 sec. The *Ss* used a cardboard strip to cover all spaces except the one in which they were writing. The series was presented twice, each time in a different random order. The interval between series was 30 sec. The *E* stated that the second presentation was for purposes of studying the variability of matches.

For each combination of materials and length there were an intentional and an incidental group. The intentional *Ss* were instructed to remember as many of the items as possible. No learning instructions were given to the incidental *Ss*.

A 5-min. test of free recall was given 30 sec. after the end of practice. The *Ss* were instructed to write down as many items as they could remember, regardless of order. They were encouraged to guess.

Subjects.—With two kinds of materials, three lengths of list, and two types of instructions, there were 12 groups of 24 *Ss* each. The *Ss* served in small groups, averaging five in number. They were assigned to the experimental conditions in rotation, with correction for equal *N's*. Two incidental *Ss* exposed to nonsense syllables (both in the 36-item group) attempted to memorize the items and were replaced. Three incidental *Ss* exposed to adjectives (two in the 36-item group and one in the 48-item group) were replaced for the same reason.

Results and Discussion

Retention of nonsense syllables.—Table 3 shows the mean number of nonsense syllables correctly reproduced. The intentional groups consistently recall more than the incidental groups. Under incidental conditions, the absolute number of correct responses increases progressively as a function of length of list. Under intentional conditions there is an increase followed by a slight decline. Thus, the increase in the absolute number of items is clearly greater under incidental than intentional conditions, and the two kinds of learners

TABLE 3

MEAN NUMBERS OF CORRECT RESPONSES AND ERRORS AS A FUNCTION OF INSTRUCTIONS AND LENGTH OF LIST

Length of List	Nonsense Syllables								Adjectives			
	Correct Responses		Intralist Confusions		Other Errors		Total Errors		Correct Responses		Errors	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
						Intentional						
24	4.83	1.77	3.83	1.91	2.29	2.02	6.12	3.05	13.46	3.29	1.96	1.89
36	5.62	2.84	4.29	3.06	1.75	1.79	6.04	3.96	16.83	4.08	1.79	2.35
48	5.29	3.09	4.96	2.57	1.50	1.12	6.46	2.84	19.38	5.95	2.17	1.95
						Incidental						
24	2.21	1.29	3.58	2.12	3.12	2.17	6.70	3.30	12.92	2.98	1.21	1.32
36	2.88	2.15	3.54	2.20	2.54	2.74	6.08	3.32	14.38	3.91	2.79	2.87
48	4.33	2.21	5.75	3.69	2.58	2.57	8.33	4.66	16.42	4.41	2.33	2.25

converge as the list is lengthened. Analysis of variance shows the overall difference between intentional and incidental learners to be significant ($F = 32.21$, $df = 1$ and 138, $P < .001$). Length of list and the interaction of length and instructions are also significant ($F = 3.38$ and 3.22, respectively; $df = 2$ and 138, $P < .05$ for both).

When retention scores are converted into percentages of the total list recalled, the functions shown in Fig. 2 are obtained. The percentages recalled by intentional learners decline as length increases, whereas the percentages recalled by incidental learners remain virtually constant. Following arcsine transformation, the percentage scores were subjected to an analysis of variance.[4] Instructions are, of course, a highly significant source of variance ($F = 33.30$, $df = 1$ and 138, $P < .001$). The significant F ratio for length of list (3.49, $df = 2$ and 138, $P < .05$) reflects almost entirely the changes under intentional conditions. The significant interaction of length of list and instructions ($F = 6.52$, $df = 2$ and 138, $P < .01$)

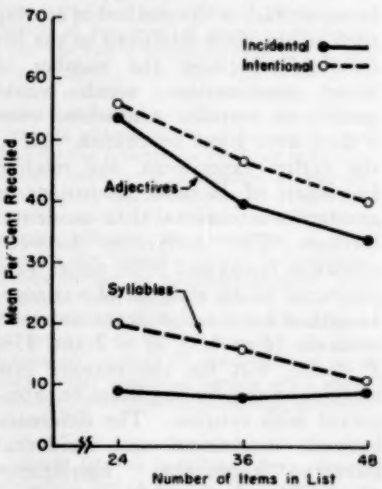

FIG. 2. Mean percentage recalled as a function of stimulus materials and length of list.

[4] The same transformation was used in the other analyses of percentage scores in this experiment.

indicates that the changes in the percentages recalled are reliably greater for the intentional than incidental groups.

Errors at recall of nonsense syllables.—The mean numbers of errors on the recall test for nonsense syllables are shown in Table 3. The frequency of errors tends to be higher for incidental learners, although there is virtually no difference for the 36-item list. The over-all difference between the two kinds of learners is not significant ($F = 1.88$, $df = 1$ and 138, $P > .05$), nor is Length ($F = 1.72$, $df = 2$ and 138, $P > .05$) or the interaction, Length × Instructions ($F = .80$). A division into intralist confusions and other errors was made on the basis of the same criteria as in Exp. I. As Table 3 shows, the relative frequency of intralist confusions increases as a function of length. This trend may reflect a rise in interference as the list is lengthened. At least some of the observed increases in confusion errors may, however, be an artifact of the method of scoring such errors, since additions to the list inevitably increase the number of letter combinations which would qualify as intralist confusions even if they were given by chance. As in the earlier experiment, the relative frequency of intralist confusions is greater for intentional than incidental learners. The differences between confusion errors and other errors were subjected to an analysis of variance. Length of list is a significant source of variance ($F = 4.90$, $df = 2$ and 138, $P < .01$), but for the reasons just mentioned this finding must be interpreted with caution. The difference between intentional and incidental learners approaches significance ($F = 3.62$, whereas 3.91 is required for significance at the .05 level with 1 and 138 df). There is no interaction of length of list with instructions ($F = .22$).

The results for the 24-item list may be compared with the 2-presentation groups of Exp. I since the same list of syllables was used. Recall is higher and errors are fewer in Exp. II in which a slower rate of presentation (10 sec.) was used than in Exp. I (4 sec.). Intralist confusions are relatively more frequent under the more difficult conditions of Exp. I. As a result, the differences in the distributions of errors are less clear-cut in Exp. II than in Exp. I. The results are consistent insofar as the incidental Ss in Exp. II again show a tendency to make a larger number of errors but relatively fewer intralist confusions.

Retention of adjectives.—The mean number of adjectives recalled is presented in Table 3. Intentional learners again recall more than incidental learners ($F = 7.13$, $df = 1$ and 138, $P < .01$). As in the case of nonsense syllables, the absolute number of items recalled increases significantly with length of list ($F = 13.68$, $df = 2$ and 138, $P < .001$). The increases in recall are, however, of comparable magnitude for the two kinds of learners, and the interaction of instructions and exercise is not significant ($F = .83$).

As Fig. 2 shows, the percentage of adjectives recalled declines as a function of length of list for both kinds of learners. In contrast to the results with nonsense syllables, the losses are no greater for intentional than incidental learners; if anything, the opposite is true. Analysis of variance of the percentage scores yields significant F ratios for Instructions (6.48, $df = 1$ and 138, $P < .01$) and for Length of List (26.59, $df = 2$ and 138, $P < .001$), but the interaction of the two variables is clearly not

significant ($F = .46$). Thus, intentional and incidental learners show parallel changes in the percentages recalled.

There are no consistent trends in the frequency of errors in the recall of adjectives. Recall is high under both conditions of learning and the number of errors is small throughout. Most of the errors appeared to be synonyms of the learning items.

Interaction of instructions, materials, and length of list.—Table 4 summarizes an over-all analysis of variance of the percentages of recall. The main effects—Materials, Instructions, and Length of List—are highly significant. The significant interaction of instructions and materials reflects the fact that the mean difference between intentional and incidental learners is greater with nonsense syllables than with adjectives. There is also an interaction of length and materials. The over-all effects of length are greater with adjectives, for which both kinds of learners show progressive drops with increases in the list; only intentional learners show this effect with nonsense syllables. The interaction, Length × Instructions, is not significant, but the significant double interaction indicates that the difference between intentional and incidental learners is a joint function of length and materials.

Thus, with nonsense syllables intentional *Ss* show considerable superiority when the list is relatively short, but this advantage decreases as the list is lengthened. These findings are consistent with the assumption that incidental *Ss* make fewer differential responses to nonsense items than to meaningful words. Hence, they retain fewer nonsense syllables than the intentional *Ss* but also are less subject to increases in intraserial interference with addition of new items to the list. With familiar meaningful items the pattern of differential responses made by the two kinds of learners becomes more similar. As a result, the differences in the amount of recall are reduced and the two conditions of learning produce parallel trends in retention scores with changes in the length of list.

TABLE 4

OVER-ALL ANALYSIS OF VARIANCE OF PERCENTAGES RECALLED IN EXP. II

Source	*df*	*MS*	*F*	*P*
Instructions	1	1620.32	33.66	.001
List length	2	1219.31	25.33	.001
Materials	1	38,438.40	798.47	.001
I × L	2	103.00	2.14	
I × M	1	209.62	4.35	.05
L × M	2	301.52	6.26	.02
I × L × M	2	213.87	4.44	.05
Error	276	48.14		

SUMMARY

This study investigated the differences between intentional and incidental learning as a function of frequency of exercise and length of list. In Exp. I frequencies of 1, 2, 4, 8, and 16 presentations were used with a list of 24 nonsense syllables. Intentional learners showed a significant upward trend in recall as a function of exercise whereas incidental learners did not. The difference between the two kinds of learners increased steadily as a function of exercise.

In Exp. II lists of 24, 36, and 48 nonsense syllables or adjectives were learned. With nonsense syllables, the percentage of recall decreases as a function of length for intentional learners but not for incidental learners. With adjectives, both kinds of learners show progressive declines in the percentage retained as the length of the list is increased. In both experiments analysis of the errors at recall indicates that intentional learners give a higher proportion of intralist confusions than incidental learners.

The results are consistent with the view that incidental learning establishes fewer correct associations and is subject to less intraserial interference than intentional learning. Hence, incidental learners benefit less from repeated presentations of the material than intentional

learners but also are less adversely affected by increases in the length of the list when the material is conducive to intraserial interference.

References

1. Brown, G. H. Factors influencing incidental learning. *J. exp. Psychol.*, 1954, **43**, 163–169.
2. Calhoun, S. W. Influence of length of lists upon ability immediately to reproduce disconnected word series auditorily presented. *J. exp. Psychol.*, 1934, **17**, 723–738.
3. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–267.
4. Haagen, C. H. Synonymity, vividness, familiarity and association value ratings of 400 pairs of common adjectives. *J. Psychol.*, 1948, **27**, 453–463.
5. Postman, L., & Adams, P. A. Studies in incidental learning: VI. Intraserial interference. *J. exp. Psychol.*, 1957, **54**, 153–167.
6. Postman, L., Adams, P. A., & Phillips, L. W. Studies in incidental learning: II. The effects of association value and of the method of testing. *J. exp Psychol.*, 1955, **49**, 1–10.
7. Saltzman, I. J., & Atkinson, R. L. Comparison of incidental and intentional learning after different numbers of stimulus presentations. *Amer. J. Psychol.*, 1954, **67**, 521–524.
8. Shurrager, H. C. The measurement of memory on an absolute scale. *Psychol. Monogr.*, 1940, **52**, No. 3 (Whole No. 234). Pp. 21–38.
9. Woodworth, R. S. The influence on retention of conditions favoring quickness of learning. *J. Philos.*, 1915, **12**, 246.

(Received September 14, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

SUPPLEMENTARY REPORT: DISCRIMINATION LEARNING IN RATS AS A FUNCTION OF HIGHLY DISTRIBUTED TRIALS

DEMPSEY F. PENNINGTON AND ROBERT THOMPSON
University of Alabama *Southeast Louisiana Hospital*

The present experiment extends a previous study (2) which reported the effects of intertrial intervals from 45 sec. to 6 min. upon acquisition of a horizontal-vertical discrimination in rats. Intervals of 40 min., 12 hr., 24 hr., and 48 hr. were used here. It was expected that the decrement in performance previously observed to occur from 3 to 6 min. would become even more marked with longer spacings between trials.

Method.—All procedural aspects of this study were the same as those described earlier (2). Male albino rats were randomly divided into four groups of 10 *S*s which were trained with the intertrial intervals specified above. Each interval began with *S*'s entry into the goal box and concluded with *S*'s placement into the start box. Upon a choice, *S* remained in the goal box for 30 sec. and was then returned to its home cage. The 24-hr. and 48-hr. groups were

TABLE 1

PERFORMANCE MEASURES TO CRITERION

Group	Trials		Errors	
	Mean	*SD*	Mean	*SD*
40 min.	26.3	13.3	12.1	5.6
12 hr.	30.4	19.9	14.8	9.4
24 hr.	34.7	15.1	16.6	8.2
48 hr.	46.2	27.2	21.0	11.1

run at noon; the 12-hr. group was run at noon and midnight; in the 40-min. group the trials were begun at noon. Training was terminated when *S* gave 9 correct responses in 10 trials or received a total of 70 trials.

Results.—Performance measures are shown in Table 1. All *S*s met the criterion within 70 trials, except one *S* in the 12-hr. group and five *S*s in the 48-hr. group. The scores of these *S*s were treated as though the criterion was reached at the point training was terminated. The data in Table 1 and the average learning curves (Fig. 1) indicate decreasing efficiency of performance with increasing degrees of trial distribution.

The trial and error measures were of heterogeneous variance. An analysis of variance (square-root transformation) yielded *F*'s of 1.86 and 1.53 for trials and errors, respectively ($P > .05$). This lack of significance is attributed largely to the failure to continue training beyond 70 trials in the 48-hr. group. Using the Mann-Whitney *U* test (1), the differences in criterion scores between the 40-min. and 48-hr. groups were significant ($P = .02$).

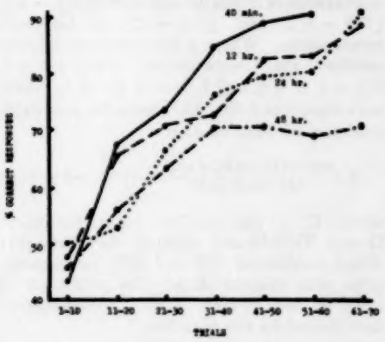

FIG. 1. Average learning curves for all groups.

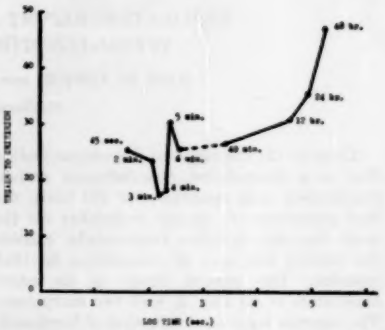

FIG. 2. Mean trials to criterion on the horizontal-vertical discrimination as a function of the intertrial interval.

Figure 2 shows the means for trials in this and the earlier study (2). It becomes apparent that the functional relationship between performance and the length of the intertrial interval is not a simple one.

Discussion.—The data suggest that increasing the intervals from 40 min. to 48 hr. has a retarding effect upon the mastery of a pattern discrimination problem. More significant, perhaps, is the finding that some rats can acquire a discrimination habit with an intertrial interval of two days. Assuming that the learning process is continuous, it is surprising that the effects of a single trial involving one nonspatial cue may be retained by rats for at least 48 hr. It is also apparent that a generalized theory of learning must incorporate some forgetting factor to account for the deteriorative effects of distribution of practice as well.

REFERENCES

1. MANN, H. B., & WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1947, **18**, 50–60.
2. THOMPSON, R., & PENNINGTON, D. F. Memory decrement produced by ECS as a function of distribution in original learning. *J. comp. physiol. Psychol.*, 1957, **50**, 401–404.

(Received July 17, 1957)

Journal of Experimental Psychology
Vol. 56, No. 1, 1958

REPLICATION REPORT: TWO- AND THREE-CHOICE VERBAL-CONDITIONING PHENOMENA[1]

JOHN W. COTTON AND ALLAN RECHTSCHAFFEN[2]

Northwestern University

Gardner (3) has reported the unique finding that in a three-choice, noncontingent verbal-conditioning task extended over 450 trials, the final percentage of response probability for the most frequent stimulus substantially exceeds the relative frequency of presentation for that stimulus. The present study is an exact duplication of his Exp. I, with two exceptions: The response boxes used here were of fiberboard, not metal, and had the stimulus light on the front, not the top. A new sequence of stimuli replaced the one Gardner developed for each group. This study includes two further replications of one group to provide added information on reproducibility of verbal-conditioning data.

Table 1 presents the means and variances for relative frequencies of "A" responses on Trials 286–450 in the present study. The identification number for each group (e.g., 70-15-15) indicates the input percentages for "A," "B," and (where applicable) "C" stimuli. No mean relative frequency in the present study differs by more than .026 from Gardner's corresponding result, and no deviation from Gardner is significant. Every mean value except in Groups 70-30 and 60-40 differed significantly from the input frequency. Two-tailed t tests with empirical rather than theoretical estimates of variance were employed with a 5% level of significance required. A four-group analysis of variance compared the present 70-15-15 groups and Gardner's 70-15-15 group, showing no significant deviations among them. Figure 1 compares the mean acquisition curves for the three replications of the 70-15-15 condition, showing close block-by-block correspondences among groups.

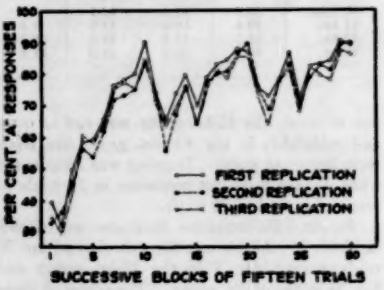

FIG. 1. A comparison of response curves for three groups given identical sequences of stimuli under Cond. 70-15-15.

The Gardner finding is contrary to predictions from statistical (2) and stochastic (1) learning theory. Gardner's own explanation, which is unsatisfactory because of indeterminate predictions in two groups, can be improved as follows: Postulate that *Ss* learn to minimize the variance of misses on each light ($[A(1-a)]$, $[B(1-b)]$, and $[C(1-c)]$ in Gardner's terminology). Where a zero variance solution satisfies the requirements $a+b+c=1$, $0 \leq a \leq 1$, $0 \leq b \leq 1$, and $0 \leq c \leq 1$, Gardner's equations hold. Otherwise the acceptable minimum variance solution is

$$a = \frac{6A^2 - 3AC + 3BC + 3AB}{6A^2 + 6AB + 6B^2},\ b = 1 - a,\text{ and } c = 0,$$

where C is the smallest input frequency. Groups 70-20-10 and 60-30-10 then have predicted a-values of .799 and .695, respectively, quite near existing data. The prediction of $c = 0$ is not verified, nor is pre-asymptotic data treated by this postulate.

TABLE 1

MEAN RELATIVE FREQUENCIES OF A RESPONSES ($p[A]$) AND VARIANCES (s_p^2) OBTAINED IN A REPLICATION OF GARDNER'S EXP. I (N = 24 per group, Trials 286–450)

Condition	$p(A)$	s_p^2
70-30	.741	.0098
70-15-15	.821	.0087
70-15-15	.809	.0104
70-15-15	.785	.0056
70-20-10	.801	.0188
60-40	.641	.0139
60-20-20	.660	.0093
60-30-10	.658	.0149

[1] This study was facilitated by funds provided by the Graduate School Committee on Research Funds, Northwestern University.

[2] Now at the University of Chicago.

REFERENCES

1. BUSH, R. R., & MOSTELLER, F. *Stochastic models for learning.* New York: Wiley, 1955.
2. ESTES, W. K. Theory of learning with constant, variable, or contingent probabilities of reinforcement. *Psychometrika*, 1957, **22**, 113–132.
3. GARDNER, R. A. Probability-learning with two and three choices. *Amer. J. Psychol.*, 1957, **70**, 174–185.

(Received May 7, 1957)

ZEITSCHRIFT FÜR EXPERIMENTELLE UND ANGEWANDTE PSYCHOLOGIE

Organ der Deutschen Gesellschaft für Psychologie

Herausgeber für Deutschland

PROF. DR. PH. LERSCH, München; PROF. DR. F. SANDER, Bonn

Herausgeber für Österreich	Herausgeber für die Schweiz	Redaktion
PROF. DR. H. ROHRACHER, Wien	PROF. DR. R. MEILI, Bern	DR. C. J. HOGREFE, Göttingen

Band IV/4

REINHARD TAUSCH, Die Frage der Lenkung und Führung in psychologisch-therapeutischen Gesprächen nicht-directiver Orientierung

GÜNTER RÖSLER, Über die Vibrationsempfindung

HEINZ HECKHAUSEN UND LILLY KEMMLER, Entstehungsbedingungen der kindlichen Selbständigkeit

WOLFGANG HORN, Zur Entwicklung, Aufgabenanalyse, Eichung und Anwendung des Begabungstestsystems (B-T-S)

THEODOR F. HAU, Vergleichende Untersuchungen an Schizophrenen und Neurotikern mit dem MMQ (Maudsley-Persönlichkeitsfragebogen)

HANS-HENNING ECKARDT, Zur Theorie der technischen Begabung

WALTER KRISTOF, Zur Frage der statistischen Sicherung von Profildifferenzen

All articles are provided with summaries in English and French.

The journal appears quarterly and contains 160 pages per number.
Subscription is DM 64.—per volume.

VERLAG FÜR PSYCHOLOGIE • DR. C. J. HOGREFE • GÖTTINGEN
(Brentanoweg 10, Germany)